

Bryn Mawr College Scholarship, Research, and Creative Work at Bryn Mawr College

Graduate School of Social Work and Social
Research Faculty Research and Scholarship

Graduate School of Social Work and Social
Research

2013

The Therapeutic Alliance and Psychotherapy Outcomes for Young Adults Aged 18 to 34: Protocol for a Systematic Review

Stacy J. Green
Bryn Mawr College

Julia H. Littell
Bryn Mawr College, jlittell@brynmawr.edu

Karianne T. Hammerstrøm

Emily Tanner-Smith



This work is licensed under a [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/).

[Let us know how access to this document benefits you.](#)

Follow this and additional works at: http://repository.brynmawr.edu/gsswsr_pubs

 Part of the [Social and Behavioral Sciences Commons](#)

Citation

Green, Stacy J.; Littell, Julia H.; Hammerstrøm, Karianne T.; and Tanner-Smith, Emily, "The Therapeutic Alliance and Psychotherapy Outcomes for Young Adults Aged 18 to 34: Protocol for a Systematic Review" (2013). *Graduate School of Social Work and Social Research Faculty Research and Scholarship*. Paper 61.
http://repository.brynmawr.edu/gsswsr_pubs/61

This paper is posted at Scholarship, Research, and Creative Work at Bryn Mawr College. http://repository.brynmawr.edu/gsswsr_pubs/61

For more information, please contact repository@brynmawr.edu.

The Therapeutic Alliance and Psychotherapy Outcomes for Young Adults Aged 18 to 34: Protocol for a Systematic Review

Stacy J. Green, Julia H. Littell, Karianne Hammerstrøm,
Emily Tanner-Smith, Jessica Schaffner Wilen

PROTOCOL



THE CAMPBELL COLLABORATION

Contact information

Lead review author:

Stacy J. Green

26 Summit Grove Avenue, Suite 206

Bryn Mawr, PA 19010

USA

S2green@brynmawr.edu

(+1) 215.803.8314

Table of contents

1	BACKGROUND	2
1.1	GENERAL	2
1.2	Description of the PROBLEM OR ISSUE	5
1.3	Description of the PHENOMENA BEING INVESTIGATED	5
1.4	HOW THE ALLIANCE MIGHT WORK	6
1.5	WHY IT IS IMPORTANT TO DO THIS REVIEW	6
2	OBJECTIVES OF THE REVIEW	7
3	METHODS	8
3.1	Criteria for considering studies for this review	8
3.2	Search methods for identification of RELEVANT studies	10
3.3	Data collection and analysis	13
4	ACKNOWLEDGEMENTS (CURRENT ONLY)	19
5	REFERENCES	20
5.1	Additional references	20
6	APPENDICES	26
6.1	Appendix 1: SCREENING AND DATA EXTRACTION FORM	26
7	CONTRIBUTION OF AUTHORS	43
8	DECLARATIONS OF INTEREST	44
9	SOURCES OF SUPPORT	45
9.1	Internal sources	45
9.2	External sources	45

1 Background

1.1 GENERAL

In recent decades, important demographic, social, and cultural changes have affected the lives and needs of young adults in many countries around the world. In high-income countries, young adults are now taking longer to finish their education, assume full time employment, and undertake personal financial responsibility (Bound, Lovenheim, & Turner, 2010; Furstenberg, Rimbaut, & Settersten, 2005; Schoeni & Ross, 2004). For example, The Organization for Economic Co-Operation and Development (OECD) shows greatly increased rates of unemployment in the “youth” (15-24) age group in countries such as Spain, Greece, and Ireland; and rising unemployment pushes many young people back home to live with their parents (Bell & Blanchflower, 2010). Based on OECD data, The Economist states that “46% of 18 to 34 year-olds in the European Union lived with at least one parent; in most countries the stay-at-homes were more likely to be unemployed than those who had moved out” (“The Jobless Young,” 2011).

Newman and Newman (2011) indicate the developmental foci for 18 to 34 year-olds are the psychosocial tasks (from Erikson, 1959) of establishing facets of identity, and constructing the hallmarks of adulthood such as intimate relationships, childbearing, and mature work. For many, these tasks now take place after age 25, and into their early 30’s. Hence, recent studies of young adulthood produced by groups such as The Transition to Adulthood (MacArthur Foundation), have focused on 18-34 year-olds. The Pew Foundation called the 25 to 34 year-old cohort, “The Boomerang Generation,” for their repeated returns to their parent’s home and financial dependence exacerbated by lower levels of economic opportunities (Parker, 2012). This preparatory period of life is putting additional strains on individuals, families of origin, and on institutions that support young adults (Berlin, Furstenberg, & Waters, 2010). At the same time, young adults are coping with marked levels of mental health issues.

Psychiatric disorders in the U.S. are most prevalent in young adulthood for ages 18 to 25, with 29.9% of that group reporting serious mental illness (U.S. Substance Abuse and Mental Health Services Administration, 2010). The U.S. General Accounting Office (GAO) estimated that “at least 2.4 million young adults aged 18 through 26—or 6.5% of the 37 million non-institutionalized young adults in that age range—had a serious mental illness in 2006” (U.S.

GAO, 2008, p. 9). The U.S. 2007 National Survey on Drug Use and Health (NSDUH) indicates that younger adults (ages 18 to 25) demonstrate higher levels of psychological distress than other age groups in the United States (U.S. Substance Abuse and Mental Health Services Administration, 2010). Kessler et al. (2007) summarized the World Health Organization's (WHO) World Mental Health Survey data on the incidence and prevalence of major mental health disorders (anxiety, mood, impulse control, and substance use) across 17 countries. Examined by age group (18 to 34, 35 to 49, 50 to 64, 65+), the WHO data show high rates of multiple types of mental disorders for 18 to 34 year olds, indicating significant levels of distress for this age group in many countries around the world (Kessler et al., 2007).

These individuals with psychological distress are at heightened risk for mental health problems due to complex and changing societal factors, and yet they are less likely than other adults to receive mental health services (10.9 among those aged 18 to 25 vs. 14.8 percent among those aged 26 to 49 and 13.6 percent among those aged 50 or older) (U.S. Substance Abuse and Mental Health Services Administration, 2010). In the Worldwide Use of Mental Health Services for Anxiety, Mood, and Substance Disorders study based on the results from 17 countries in the WHO World Mental Health (WMH) Surveys (Kessler et al., 2007) the authors state, "age was a significant predictor of receiving mental health services in eight countries; in these, respondents in the middle years of life were generally more likely to receive services than either those younger or older" (Wang et al., 2007, pg. 6). However, compared with the voluminous literature on adults, fewer studies and no systematic reviews examine the effects of the alliance on psychotherapeutic outcomes specifically for young adult clients.

Among the elements that may lead to successful psychotherapy, the therapeutic alliance is considered both a central and a common factor linked to outcomes in psychotherapy by many leading psychotherapists (Bordin, 1979; Freud, 1912/1958; Rogers, 1951) as well as proponents of the "Common Factors" orientation (Rosenzweig, 1936; Wampold, 2001). "Common Factors" refers to the perspective that effectiveness in psychotherapy is due to factors that are common to all forms of therapy rather than to specific techniques. Although considered integral to the process of psychotherapy, leading researchers have disagreed about exactly what the alliance is and how it works. Authors have used diverse theoretical conceptualizations to describe (Bordin, 1983; Freud 1912/1958; Greenson, 1965; Zetzel, 1956), operationalize, and measure the alliance (Gaston & Marmar, 1994; Horvath, 1981; Luborsky, Crits-Cristophe, Alexander, Margolis, & Cohen, 1983; Marmar, Horowitz, Weiss, & Marziali, 1986; Suh, Strupp, & O'Malley, 1986).

The variety of therapeutic alliance constructs is reflected in the diversity of measurement instruments that have arisen out of different psychotherapeutic frameworks. For example, the Working Alliance Inventory (Horvath & Greenberg, 1986) is a 36-item scale with three subscales reflecting Bordin's (1983) pantheoretical, tripartite conceptualization of the working alliance, which includes agreement on the goals of treatment, agreement on the tasks of treatment, and affective bonds. The Penn Helping Alliance, a shorter 10-item scale,

evaluates two separate aspects of the alliance called Type I and Type II. Type I rates the client's feelings about the therapist from a psychoanalytically-influenced perspective and Type II measures therapist and client agreement on the tasks and goals of treatment, similar to the Working Alliance Inventory and Bordin's conceptualization (Fenton, Cecero, Nich, Frankforter, & Carroll, 2001).

In their meta-analysis, Horvath et al. (2011) noted that there are at least 30 different alliance measures and multiple studies show the leading measures to be related to each other (Bachelor, 1991; Horvath & Luborsky, 1993; Tichenor & Hill, 1989). In their factor analysis of the patient version of three leading measures, Hatcher and Barends (1996, p. 1328) found "the total scores on the three measures correlated highly: CALPAS and WAI, $r = .85$; CALPAS, and HAQ, $r = .74$; WAI and HAQ, $r = .74$ ($p < .0001$, $N = 231$), indicating the presence of a strong general factor." However, the authors state "there has been little evidence to support the theoretical dimensions that underlie the measures." Their analysis of the Working Alliance Inventory, California Psychotherapy Alliance Scales, and the Helping Alliance Questionnaire indicated that after removing the large general factor, only two of the six factors identified using principal component analysis, "Confident Collaboration and Idealized Relationship, correlated with patients' estimate of improvement ($r_s = .37$ and $-.23$, respectively; $p < .001$)" (Hatcher & Barends, 1996, p. 1326).

Given the large body of empirical research on the therapeutic alliance and its relation to outcomes, meta-analyses have focused on specific populations such as adults (18+) (Horvath & Bedi, 2002; Horvath & Symonds, 1991; Horvath et al., 2011; Martin et al., 2000), youth (under 19) (McLeod, 2011; Shirk, Karver, & Brown, 2011; Shirk & Karver, 2003), youth and families (Karver, Handelsman, Fields, & Bickman, 2006), and couples and families (Friedlander, Escudero, Heatherington, & Diamond, 2011). We know of no meta-analyses on the alliance and outcomes specifically looking at young adults.

Between 1991 and 2001 four major meta-analyses examined relationships between the therapeutic alliance and psychotherapy outcomes for adults in individual psychotherapy (Horvath & Bedi, 2002; Horvath & Symonds, 1991; Horvath et al., 2011; Martin et al., 2000). These meta-analyses produced similar, moderate correlational effect sizes (ES) ranging from $r = .21$ (Horvath & Bedi, 2002) to $r = .28$ (Horvath, Del Re, Fluckinger, & Symonds, 2011).

Four additional meta-analyses examined correlations between the alliance and outcomes in psychotherapy for youth (children and adolescents) in individual and family treatment (Shirk & Karver, 2003; Karver et al., 2006; McLeod, 2011; Shirk, Karver, & Brown, 2011). These studies found correlational ESs of $r = .24$, $r = .17$, $r = .14$, and $r = .22$ respectively. Two authors noted methodological issues, such as study heterogeneity, might have impacted their effect sizes.

We believe that all of the previous meta-analyses of research on the alliance and outcomes for adult clients in individual psychotherapy have limitations when compared to currently recommended methodological standards. The AMSTAR instrument, developed to assess the

methodological quality of systematic reviews and meta-analyses (Shea et al., 2007), contains 11 items identified by exploratory factor analysis performed on over 150 studies to identify core components of review quality. Using the AMSTAR criteria, the four meta-analyses performed on studies of adults (Horvath & Bedi, 2002; Horvath & Symonds, 1991; Horvath, et al., 2011; Martin et al., 2000) appear to lack many of elements of rigorous and valid research syntheses. For example, none of the four meta-analyses reported that they had a public, a priori design, duplicate study selection and data extraction, or formal evaluation of study quality.

1.2 DESCRIPTION OF THE PROBLEM OR ISSUE

The hazards of psychological setbacks to young adults entering employment, family commitments, and higher education can be severe: mental health issues increase risk for many long-term, negative consequences such as not finishing education, unplanned pregnancy, drug abuse, and unemployment (Gralinski-Bakker Hauser, Billings, & Allen, 2005). In addition, young adulthood is a period when some individuals are experiencing their first episodes of mental illness and are especially challenged (Pottick et al., 2008). Yet, two studies based on nationally representative US samples (Kessler et al., 2005; Pottick et al., 2008) indicate that individuals aged 18 – 24 were “significantly less likely to receive mental health services than adults in older age groups” (Pottick et al., 2008, p. 382). Pottick et al. (2008) indicate that multiple forces may limit utilization for this age group including loss of health insurance, managed care, challenges to continuity of care, and changes in the mental health needs of this cohort. For example, continuity of care is often challenging for individuals transitioning from child mental health outpatient and residential services from which they are ‘aging out’ to fewer available adult services (Pottick et al., 2008). Some evidence indicates that the prevalence and seriousness of mental health problems may be increasing among young adults, as college counseling center directors report growing numbers of students with more complex and severe mental health problems (Gallagher, 2010).

1.3 DESCRIPTION OF THE PHENOMENA BEING INVESTIGATED

This systematic review will evaluate the relationship between the therapeutic alliance and psychotherapy outcomes, using observational studies of young adults (ages 18-34) that include reliable and valid measures of the therapeutic alliance and evaluate psychotherapy outcomes. The strength of the alliance is expected to predict more positive outcomes.

1.4 HOW THE ALLIANCE MIGHT WORK

Many studies identify the quality of the therapeutic alliance as one of the major factors that may lead to positive outcomes across many forms of psychotherapy (Glencavage & Norcross, 1990; Weinberger, 1995). Because the process of relationship formation occurs within most, if not all, therapeutic modalities, the alliance has come to be considered a "common factor" or universal component of successful therapy (Rosenzweig, 1936, Frank & Frank, 1991, Wampold, 2001). Theoreticians from Freud onward have pointed to various aspects of the therapeutic relationship that may result in positive outcomes for the client. Freud (1912) hypothesized that the client's transference or prior memories and experiences could imbue therapist with both "authority" and credibility and lead to therapeutic efficacy. This theory led to an on-going debate that divides the relationship into elements from the past (commonly known as the transference) and the elements that are created in real time between the client and therapist. Rogers (1957) elevated the importance of the therapeutic alliance when he identified therapist-offered "necessary and sufficient" conditions including genuineness, empathy, and unconditional positive regard as responsible for therapeutic personality change. More recent theory stresses a pantheoretical and generic orientation with three major elements: (1) agreement and shared understanding of the goals of therapy, (2) agreement on tasks to be undertaken during therapy, and (3) an affective bond between therapist and client (Bordin, 1983). Together, these factors are thought to be supportive of the client's progress as he or she has defined it.

1.5 WHY IT IS IMPORTANT TO DO THIS REVIEW

We know of no systematic review or meta-analysis using current methodology that examines the association between the therapeutic alliance and the outcomes of psychotherapy specifically for young adults. There is potential for young adults to have unique psychotherapeutic needs based upon developmental challenges associated with the transition to adulthood and recent demographic and economic changes that have altered the social landscape and opportunities for young adults (Berlin et al., 2010). In some countries, many of the developmental tasks that used to be achieved earlier (financial independence, marriage, childbearing etc.) have now been delayed to the latter part of the age range, placing additional stresses on individuals who have high levels of psychological needs (Bound, Lovenheim, & Turner, 2010; Furstenberg, Rimbaut, & Settersten, 2005; Kessler et al., 2005; Schoeni & Ross, 2004). Using recent advances in meta-analysis methodology, this systematic review may be able to provide important insights into the role of the therapeutic alliance in psychological treatment of this important population.

2 Objectives of the review

The objective of this review is to summarize the available literature and produce reliable estimates of associations between measures of the therapeutic alliance and outcomes of psychotherapy for 18 to 34 year olds.

Where possible, we will:

- Determine whether client, therapist, or observer reports of the alliance are better predictors of outcomes.
- Analyze the level of agreement between client and therapist ratings of the alliance as a possible predictor of outcomes.
- Evaluate moderators such as age, alliance measure, outcome, alliance rater, and time of alliance assessment that may influence the relationship between the alliance and psychotherapy outcomes.
- Address issues of time order and causality by synthesizing results from studies that use baseline characteristics and/or scores on initial outcome measures as control variables in analyses of associations between later alliance measures and more distal outcome scores.

3 Methods

3.1 CRITERIA FOR CONSIDERING STUDIES FOR THIS REVIEW

3.1.1 Types of studies

This review will include observational, cohort studies of the therapeutic alliance and psychotherapy outcomes for young adults in face-to-face psychotherapy with mental health professionals. Eligible studies must include at least one measure of the therapeutic alliance that preceded at least one outcome measure. This is necessary to establish time order (the alliance precedes outcome), one of several conditions that need to be met to support causal inferences. Cross-sectional studies in which all alliance and outcome measures were obtained at the same point in time will be excluded. Studies that attempt to manipulate or influence the alliance itself (studies in which the alliance is an outcome) will be included. Since the quality of the relationship between therapist and client is not something that can be assigned at random, randomized controlled trials, controlled before-and-after designs, and interrupted time series studies are irrelevant. Qualitative studies will be excluded.

3.1.2 Types of participants

Study participants must be young adults (ages 18-34). Studies that include wider age groups will be included only if they report alliance-outcome associations specifically for young adults as defined here. There will be no exclusions based on diagnosis, treatment modality, or severity of illness as we know of no evidence that indicates that any of these factors necessarily negate the formation of an alliance. This means we will include a wide range of diagnoses such as depression, anxiety, personality disorders, and substance abuse disorders.

3.1.3 Types of interventions

We will include all bona fide psychotherapies delivered to individuals in face-to-face interactions with mental health professionals or trainees. Interventions must be performed by mental health professionals or those in psychotherapy training programs supervised by Master's level or doctoral level mental health professionals. Activities outside of formal mental health services (such as guidance counselling, mentoring, peer counselling) are excluded. Interventions delivered by paraprofessionals are also excluded, because mental health professions require licensure or supervision of students in training programs leading

to licensure. Therapy must be provided to individuals in face-to-face sessions. We will exclude studies in which therapy was performed using online technologies, because the role of the therapist is often unclear in these interventions. We will exclude couples, family, and group therapies because alliances are much more complex and difficult to measure in these contexts than in individual treatment. All types of bona fide psychotherapies are included, since the alliance may be central to all. Examples of bona fide therapies include Psychodynamic, Emotion-Focused Therapy, Gestalt, Cognitive Behavioural Therapy, and Behavioural Therapy. Any duration of psychotherapy is acceptable since some studies show that psychotherapy can be effective in as little as one session (Borsari & Carey, 2000; Raue, Castonguay, & Goldfried, 1993).

3.1.4 Types of outcomes

The primary outcomes in this review are ratings of mood disorders such as depressive symptoms (measured, for example, by the Beck Depression Inventory; Beck et al., 1961 or the Hamilton Rating Scale for Depression; Hamilton, 1960), psychiatric symptoms (measured, for example, by the Symptom Checklist-90-R; Derogatis, 1996), global functioning (measured, for example, by the Global Assessment of Functioning (GAS; Endicott et al., 1976) or assessments of change), specific outcomes (e.g., target complaints or estimations of drug use), and treatment participation or termination status (e.g., whether the client attended the agreed upon number of sessions or terminated early) (Martin et al., 2000). Studies may include multiple measures to evaluate outcomes in more than one category of symptoms or functioning. We will only include outcome measures for which there is published evidence of reliability and validity, and we will determine whether this information was obtained in a sample comparable to the one in the included study. This requirement will result in the exclusion of purely clinical assessments (or opinions) with no known reliability or validity, as these could introduce additional bias or error into our analysis.

This review has no secondary outcomes.

3.1.5 Types of data

Alliance measures are primarily reported in the format of continuous scales. Psychotherapy outcome measures can be continuous (e.g., degree of symptom reduction, days of hospitalization) or dichotomous (e.g., any hospitalization, relapse). We expect that most of the data in this review will be in the form of correlations between alliance and outcome measures. These associations are usually expressed in terms of the Pearson product-moment correlation coefficient (r) and similar metrics. If we find studies that controlled for other variables in analyses of associations between alliance and outcome measures, we will explore options for using partial correlation coefficients or regression coefficients in our analysis.

Data may have been obtained from direct observation, interviews, self-reports, and/or clinical records. As indicated above, over 30 alliance measures have been used to measure

the alliance; however, we will only include studies that used alliance and outcome measures that have some published empirical evidence of reliability and validity in the sample studied or within a comparable sample. Elvins and Green (2007), for example, produced a comprehensive description of 32 leading alliance measures and related theoretical frameworks. For each measure, the authors listed any published evidence of reliability or validity (e.g., concurrent or criterion validity, internal consistency assessed with Chronbach's alpha, inter-rater reliability, etc.), the sources of this information, and comments about the relative strengths and weaknesses of each measure. Multiple compendiums, such as the Mental Measurements Yearbook (and others), list similar information for published outcome measures.

3.2 SEARCH METHODS FOR IDENTIFICATION OF RELEVANT STUDIES

A comprehensive search strategy may uncover studies that were overlooked in previous meta-analyses, by paying particular attention to the inclusion of grey literature. The search process pits the likelihood of identifying too many unrelated titles against the risk of missing meaningful titles (Hammerstrøm, Wade, & Jørgensen, 2010). The search will include multiple resources including electronic sources, references lists in previous meta-analyses, and citations in eligible research.

There are no language restrictions. Translation will be acquired as resources permit.

3.2.1 Electronic searches

The following databases will be searched:

AMED (Allied and Complementary Medicine)

Campbell Library

CINAHL

The Cochrane Library

Dissertation Abstracts International

EMBASE

ERIC

LILACS

MEDLINE

OpenGrey

ProQuest

PsycINFO

Sociological Abstracts

Social Services Abstracts

SveMed+

Web of Science®

3.2.2 Search terms

Search terms for OVID PsycINFO (modified as necessary for other databases) will be as follows:

1	therapeutic alliance/
2	alliance.tw.
3	((psychotherap* or psycho-therap* or psychologist* or therap* or helping or working or social) adj1 (bond* or relation*)).tw.
4	((therapist-patient* or psychologist-patient* or psychotherapist-patient* or psycho-therapist-patient* or professional-patient* or therapist-client* or psychologist-client* or psychotherapist-client* or psycho-therapist-client* or professional-client*) adj1 (bond* or relation*)).tw.
5	or/1-4
6	exp Psychotherapy/
7	(psychotherap* or psycho-therap* or psychoeducat* or psycho-educat* or psychodynam* or psycho-dynam* or psychoanaly* or psycho-analy*).tw.
8	((group or behavio* or cogniti* or general or network* or social or supporti* or interpersonal or individual or family or families or brief or psycho* or vocation*) adj2 (therap* or counsel* or rehabilitat* or habilitat* or support* or intervention*

	or train*).tw.
9	6 or 7 or 8
10	5 and 9

3.2.3 Searching other resources

As detailed below, we will conduct hand-searches, grey literature searches, website searches, and reference harvesting and branching (searching references of known studies to find additional material). We will contact authors of included studies and experts in the field to request copies of potentially eligible studies that are unpublished.

Hand search

Three journals relevant to alliance studies will be hand-searched to locate studies not identified from database searches. We will also evaluate reference lists of alliance studies and published reviews to identify additional studies. If articles are considered relevant they will be retrieved and reviewed to see if they meet the stated inclusion criteria.

To identify journals to be hand searched, we reviewed the studies included in the most recent meta-analysis of the therapeutic alliance for adults (Horvath, 2011) and identified the journals most likely to contain relevant material on the therapeutic alliance. The results of the review indicate that 1) *Journal of Consulting and Clinical Psychology*; 2) *Psychotherapy Research*; and 3) *Psychotherapy* warrant hand searching for the timeframe from 1980 forward.

Grey literature

In addition to the previous strategies, this review will include searches to locate relevant grey literature. We will search: 1) Conference proceedings from The American Psychological Association and Society for Social Work and Research; 2) Government reports in the US from National Institute of Mental Health; and 3) Professional organizations' web sites: www.naswdc.org, www.apa.org, and www.bps.org.uk.

We will search grey literature websites for relevant unpublished material: Grey.Net (<http://www.greynet.org/index.html>) and GrayLit Network (<http://graylit.osti.gov/>), *Google Scholar* (www.scholar.google.com) and the Information for Practice website (ifp.nyu.edu).

Personal communication

We will communicate with leading authors and experts in the fields of the therapeutic alliance and young adults through personal discussions, emails, and requests for information. We will ask alliance and young adult authors, practitioners, and academics for their help locating additional material that may have been missed in our searches. The alliance experts include: Jacques Barber, PhD, University of Pennsylvania; Louis Castonguay, PhD, Penn State University; Adam Horvath, PhD, Simon Fraser University; Marc Karver, PhD, University of South Florida; and Bruce Wampold, PhD, University of Wisconsin. Some of the young adult authors are Jeffrey J. Arnett, PhD, Clark University; Frank F. Furstenberg, PhD, University of Pennsylvania; and Richard A. Settersten, PhD, Oregon State University. These informants will be used to generate a snowball sample of others whom we will ask about additional relevant contacts.

3.3 DATA COLLECTION AND ANALYSIS

3.3.1 Description of methods used in component studies

The studies in the therapeutic alliance literature are observational, but vary according to research design, treatments, perspective of the person rating the therapeutic alliance (client, therapist, and/or observer), alliance measure, and outcome measures used. One study that meets our inclusion criteria is Baldwin, Wampold and Imel (2007) based on the sample's age group (18-34), types of interventions (performed by licensed mental health therapists), alliance and outcome measures with reported reliability and validity, timing of measures (at least one alliance measure before outcome measure), and type of reported data (Pearson's correlation). This study evaluated the relationship between the therapeutic alliance and psychotherapy outcome using patient outcomes from the database of the Research Consortium of Counseling and Psychological Services in Higher Education (USA). The authors analyzed data from 331 clients seen by 80 therapists. This sample only included patients who had completed the Working Alliance Indicator prior to session four and had attended at least four therapy sessions. Outcomes were analyzed using a completed initial and final Outcome Questionnaire-45. In addition to a Pearson's correlation for alliance to outcome, this study also attempted to explore the contribution of patient and therapist variability to outcomes. The Pearson's correlation produced for this study was $r = .24$ indicating that better alliances were associated with better outcomes for study participants.

In the most recent meta-analysis of alliance-outcome studies, Horvath et al. (2011) employed a number of statistical procedures to calculate r for primary studies that had reported data in various other formats. We cannot replicate these calculations based solely on published information; hence, we will need additional information from authors in order to calculate r accurately and transparently for some primary studies.

3.3.2 Criteria for determination of independent findings

We expect to encounter dependent effect sizes in most studies. These are likely to include repeated measures (over time) and multiple alliance measures (scales and subscales). There is also the likelihood that multiple clients have seen the same therapist. It is also possible that the same researchers conducted multiple studies. Because we expect to find several types of dependencies in this literature, we will estimate robust standard errors (Hedges, Tipton, & Johnson 2010) in all analyses (as outlined in detail below). In addition, we may identify multiple reports published from individual studies and may need to contact authors if we cannot clarify whether findings are independent.

3.3.3 Details of study coding categories / data extraction

Prior to screening, results of multiple search strategies will be merged into one list of titles and abstracts, deleting any duplicates. If the search yields more than 10,000 citations, we will explore the use of semi-automated screening (Wallace et al., 2010) and other strategies to expedite the screening process. When necessary, two reviewers will independently read and evaluate titles and abstracts. We will obtain the full text for any report that may meet the eligibility criteria. We will link together any reports that originated from the same study to eliminate the possibility of counting single studies multiple times. Two review authors will independently read full text reports and will then compare notes, creating a consensus list of eligible studies. If missing data prevent us from being able to decide upon a study's eligibility we will correspond with the researchers to obtain the data to facilitate a decision. The review authors will not be blinded to identifying information on journals, authors, affiliations or outcomes. If the review authors disagree about a study's eligibility, they will work together to resolve the issue with possible input from a third evaluator who has methodological proficiency. The final outcome will be a list of both included and excluded studies with a brief rationale for each exclusion decision. In addition, we will produce a PRISMA flow chart, documenting the status of studies and decision points (Moher et al., 2009).

Two review authors will separately evaluate, extract and document the pertinent study information using data extraction forms shown in Appendix 1. Initially, the extraction forms will be tested on two studies and any needed adjustments will be made before beginning the project. Two review authors will perform the study data extraction and then meet and compare forms. Disagreements will be resolved by adjusting the coding or forms, if necessary. Any changes in data extraction forms will be applied retroactively to all studies that have already been coded. Should any issue be unresolved in the data extraction process, the review authors may contact the primary study authors for more information. If this fails to enable resolution, we will consult a third review author. To evaluate reliability we will compute Kappa coefficients on all items involved in eligibility decisions and risk of bias assessment on coding forms 2 and 5.

Five forms have been designed to record data extracted from the reports evaluated in this study. Form 1 is used for initial screening to determine if the study meets basic inclusion

criteria such as subject matter, outcomes, and study design. Form 2 is used to evaluate eligibility criteria details such as alliance and outcome measures, and population. Form 3 records study level data extraction criteria such as sampling, recruitment, and setting elements. We will extract data on the location of each study, including city/region, state/province, and country. Form 4 captures reported outcome data and data collection methods. Form 5 reports study quality standards including evaluation of multiple forms of bias and potential conflicts of interest. The forms are located in Appendix 1.

3.3.4 Dealing with missing data

If we identify missing data in a report, we will contact the authors of the study to attempt to obtain data. Information on therapy dropout and attrition from research will be recorded for each study. We will assess attrition (missing cases) and the potential for outcome reporting bias (missing data) as part of our evaluation of study quality (risk of bias). If we have sufficient studies and we observe high levels of attrition in some studies, we will conduct sensitivity analyses to assess the impact of excluding studies with high levels of attrition. Similarly, we will explore for possible associations between risk of selective outcome reporting and results.

3.3.5 Assessment of risk of bias in included studies

Study reports will be evaluated for potential sources of bias associated with: selection, allocation procedures (allocation to therapist), performance, attrition, detection, unstandardized observation periods, unstandardized alliance measures, unstandardized outcome measures, conflicts of interest, and selective outcome reporting (see Appendix 1, Level 5). Further analysis will determine whether all outcomes are reported as well as whether all individuals in the sample are accounted for in the analysis.

The GRADE approach will be used to analyze the quality of evidence related to each of the key outcomes (GRADE Working Group, 2004; Guyatt, Oxman, Vist, Kunz, Falck-Ytter, Alonso-Coello, & Schünemann, 2008). The potential ratings for each outcome are High, Moderate, Low, or Very Low.

3.3.6 Data synthesis

We expect effect sizes to be correlated within studies because each study will likely provide multiple estimates of associations between alliance and outcome measures. The covariance structure of these effect sizes is not likely to be reported in any of the studies (Hedges & Olkin, 1985), so we will use a newly developed technique to handle statistically dependent effect sizes through the estimation of robust standard errors (RSE) (Hedges, Tipton, & Johnson 2010). The RSE approach estimates standard errors based on empirical assumptions of the variance, reducing the need for assumptions about the distribution of the effect size estimates. The few existing assumptions of the RSE approach are usually met; simulation studies demonstrate that confidence intervals and p-values in samples as small as

10 studies (to estimate average effect size) or between 20-40 studies (to estimate slope) approach the correct size (Hedges et al., 2010). This method therefore allows multiple effect size estimates from the same study to be included in a meta-analysis. The RSE method is an improvement upon other methods such as selecting one effect size per study, or creating a synthetic average effect size for each study, which throw away potentially valuable information about and their variability within a study. If we find sufficient studies, we plan to use RSE to model all of the types of outcomes within the same analysis (mood, symptoms, treatment completion, etc.) rather than using separate models.

Due to the expected variety of alliance and outcome measures, we plan to use a random effects model for all analyses that account for both within study sampling variance and between study variability (τ^2). Estimating τ^2 within the RSE approach requires an estimate of the correlation between all the pairs of effect sizes (ρ). Because ρ is rarely reported in primary studies, as recommended by Hedges et al. (2010) we will use a sensitivity approach to estimating τ^2 by estimating values of τ^2 at different values of ρ ranging from 0 to 1.0. Presuming there is little variation in the of τ^2 estimates at different values of ρ , we will use a fairly conservative value of $\rho = .80$ for the final τ^2 estimate to be used in all models.

As recommended by Hedges, Tipton, and Johnson (2010), we will use a conservative approach when estimating the random effects weights, and assume $\rho = 1.00$ for these calculations. In this approach, each effect size within a study is assigned a random effects weight equal to the inverse of the number of estimates in that study multiplied by that study's variance (i.e., the sum of the mean within study variance plus the between studies variance component). This conservative approach means that any given study does not receive additional weight simply because it provides multiple effect sizes. The software we will use in this review will include Microsoft Word, Microsoft Excel, and Stata. The graphical displays that will be used are Forest plots and Contour Enhanced Funnel plots (Palmer et al., 2008).

Since there is a risk of a non-normal distribution using Pearson's r , we will use Fisher's z transformation, which converts Pearson's r to the normally distributed variable z .

3.3.7 Assessment of heterogeneity

Heterogeneity of results across studies will be assessed informally by checking for overlapping confidence intervals. Formal tests of heterogeneity will be based on the Chi-square (Q) test and the I^2 statistic (Higgins & Thompson, 2002). The interpretation of the Chi-squared test will be conditional due to its low statistical power. We expect some unexplained heterogeneity and will therefore use a random-effects meta-analysis model.

3.3.8 Moderator analysis

If possible, we will use meta-regression analysis with RSE (Hedges, Tipton, & Johnson 2010) to examine factors associated with the strength of alliance-outcome associations. The potential effect size moderators of interest are the age (e.g., younger 18-24, older 25-34), alliance measure (e.g., Helping Alliance Scales (HAS), Treatment Alliance Scales (TAS), Working Alliance Indicator (WAI), alliance rater (therapist, client, observer), time of alliance assessment (e.g., early, middle, late), outcome and outcome measure (e.g., Beck Depression Inventory, SCL-90-R (Derogatis, 1996), GAS), type of treatment (e.g., Psychodynamic, Rogerian, Cognitive Behavioral), pre-therapy severity of impairment (e.g., mild, moderate, severe), number of sessions, and experience of therapist (e.g., less than three years, more than three years).

One advantage of the RSE approach is that the inclusion of multiple effect sizes per study permits examination of within-study effects of moderators that vary within studies as well as between studies. Four of the moderators of interest may vary within and between studies: rater, timing, type of alliance measure, and type of outcome measure. If any of these moderators do vary within and between studies, we will create two variables for each moderator: (1) the mean value of the moderator within each study used to estimate the between-study effect of the moderator, and (2) a study-mean centered value of that moderator to estimate the within-study effect of the moderator). The remaining four moderators (type of treatment, pre-therapy severity of impairment, number of sessions and experience of therapist) are expected to only vary between studies, and will be included as standard moderators in the meta-regression model. We will also examine and present the bivariate correlations between all moderators to assess any potential confounding among them.

3.3.9 Sensitivity Analysis

We will conduct several types of sensitivity analyses. First, we will conduct sensitivity tests with different values of ρ in the RSE models, to examine whether this assumed correlation value had any impact on the estimate of τ^2 and the resulting parameter estimates and standard errors. Second, we will examine the potential impact of missing data on the review's findings. We will also remove studies with high risk of bias of specific types (e.g., high risk of selective reporting) to see whether an analysis of the studies with low risk of bias gives a different result. Finally, we will use the Egger test (Egger et al., 1997), trim and fill analysis, and Contour Enhanced Funnel plots (Palmer et al., 2008) to check for the possibility of publication bias and small sample bias.

3.3.10 Roles and responsibilities

Stacy J. Green will lead the writing of the protocol, including the background sections. She will compile the search terms and perform some of the searches, review documents for inclusion and exclusion criteria, extract and compile data, conduct the analysis and synthesis of data, and lead the writing of the review.

Julia H. Littell will review and edit the protocol, the pilot testing, and the review as it progresses, providing substantive and methodological guidance and direction as needed.

Karianne Hammerstrøm will design and carry out the search strategy, performing many of the database searches. She will also write and/or edit the sections in the protocol that refer to the search and will also write the search sections of the review.

Emily Tanner-Smith will write the protocol sections on data synthesis. She will provide consultation on the synthesis and will review and edit the data synthesis sections of the review including discussion of unit of analysis issues.

Jessica Schaffner Wilen will review and edit the protocol as it progresses; review documents for inclusion and exclusion criteria; and extract, compile and review data.

3.3.11 Preliminary timeframe for completion of review

After the protocol is approved, we estimate that the search will take approximately six months. Following the initial search, coding and data extraction will take two to three months and analysis will take three months. Therefore, the timeframe for the entire review is one year in total.

3.3.12 Plans for updating the review

Stacy J. Green will be responsible for updating this review every three years in accordance with Campbell Collaboration guidelines and as funding becomes available.

4 Acknowledgements (*current only*)

This study is supported by the Rivitz Fellowship awarded to Stacy J. Green by the Bryn Mawr College Graduate School of Social Work and Social Research.

5 References

5.1 ADDITIONAL REFERENCES

References

- Bachelor, A. (1991). Comparison and relationship to outcome of diverse dimensions of the helping alliance as seen by client and therapist. *Psychotherapy: Theory, Research, Practice, Training*, 28(4), 534-549.
- Baldwin, S. A., Wampold, B. E. & Imel, Z. E. (2007). Untangling the alliance–outcome correlation: Exploring the relative importance of therapist and patient variability in the alliance, *Journal of Consulting and Clinical Psychology*, 75(6), 842–852.
- Beck, A.T., Ward, C. H., Mendelson, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. *Archives of General Psychiatry*, 4, 561-571.
- Bell, D. N., & Blanchflower, D. G. (2010). Recession and unemployment in the OECD. In *CESifo Forum* (Vol. 1, p. 2010).
- Berlin, G.L., Furstenberg, F.F., & Waters, M.C. (2010). Introducing the issue. *The Future of Children*, 20(1), 3-18.
- Bordin, E.S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, research & practice*, 16(3), 252-260.
- Bordin, E.S. (1983). A working alliance based model of supervision, *The Counseling Psychologist*, 11, 35-42.
- Borsari, B., & Carey, K. B. (2000). Effects of a brief motivational intervention with college student drinkers. *Journal of Consulting and Clinical Psychology*, 68(4), 728-733.
- Bound, J., Lovenheim, M.F., & Turner, S. (2010). Increasing time to baccalaureate degree in the United States, Cambridge, MA: *National Bureau of Economic Research* (NBER Working Paper No. 15892).
- Derogatis, L. R. (1996). *SCL-90-R: Symptom Checklist-90-R: Administration, scoring, and procedures manual*. NCS Pearson.

- Egger, M., Smith, G.D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal* 315, 629-634.
- Elvins, R. & Green, J. (2007). The conceptualization and measurement of therapeutic alliance: An empirical review, *Clinical Psychology Review* 28 (2008) 1167–1187.
- Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The Global Assessment Scale: a procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry*, 33(6), 766.
- Erikson, E. H. (1959). Identity and the life cycle: Selected papers. *Psychological Issues; Psychological Issues*.
- Fenton, L.R., Cecero, J.J., Nich, C., Frankforter, T.L., & Carroll, K.M. (2001). Perspective is everything: the predictive validity of six working alliance instruments, *The Journal of Psychotherapy Practice and Research*, 10(4), 262-268.
- Frank, J.D., & Frank, J.B. (1991). *Persuasion and healing: A comparative study of psychotherapy* (3rd ed.). Baltimore, MD: Johns Hopkins University Press.
- Friedlander, M.L., Escudero, V., Heatherington, L., & Diamond, G. (2011). Alliance in couple and family therapy, *Psychotherapy*, 48 (1), 25-33.
- Freud, S. (1912/1958). The dynamics of transference [Zur Dynamik der Übertragung]. (J. Starchey, Trans.). In J. Starchey (Ed.), *The standard edition of the complete psychological works of Sigmund Freud* (vol. 12, pp. 99–108). London: Hogarth Press.
- Furstenberg, F.F., Rimbaut, R.G., & Settersten, R.A. (2005). On the frontier of adulthood. In Settersten, Jr., R.A, Furstenberg, Jr., F.F., & Rumbaut, R.G., (Eds.), *On the Frontier of Adulthood: Theory, Research, and Public Policy*. Chicago: University of Chicago Press.
- Gallagher, R.P. (2010). *The National Survey of College Counseling Center Directors*. The International Association of Counseling Services, Inc.: Alexandria, VA.
- Gaston, L. & Marmar, C. (1994). The California Psychotherapy Alliance Scales. In A. O. Horvath & L.S. Greenberg (Eds.) *The Working Alliance: Theory, Research and Practice*. New York: Wiley.
- Glencavage, L., & Norcross, J. (1990). What are the commonalities among the therapeutic factors. *Professional Psychology: Research and Practice*, 21, 372-378.
- GRADE Working Group (2004). Grading quality of evidence and strength of recommendations. *BMJ*, 328, 1490-1494.
- Gralinski-Bakker J.H., Hauser, S.T., Billings, R.L. & Allen, J.P. (2005). "Risks along the Road to Adulthood: Challenges Faced by Youth with Serious Mental Disorders," in D.W.

- Osgood, E.M. Foster, C. Flanagan, & G.R. Ruth (Eds.) *On Your Own without a Net* (pp. 272–303). Chicago: University of Chicago Press.
- Greenson, R.R. (1965). The working alliance and the transference neuroses. *Psychoanalysis Quarterly*, 34, 155–181.
- Guyatt, G.H., Oxman, A.D., Vist, G.E., Kunz, R., Falck-Ytter, Y., Alonso-Coello, & Schünemann, H.J. (2008). GRADE: An emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 336, 924-926.
- Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery, and Psychiatry*, 23(1), 56.
- Hammerstrøm, K., Wade, A., & Jørgensen, A.M.K. (2010). Searching for studies: A guide to information retrieval for Campbell Systematic Reviews. *Campbell Systematic Reviews 2010: Supplement 1*, DOI: 10.4073/csrs.2010.1.
- Hatcher, R.L., & Barends, A.M., (1996). Patient's view of the alliance in psychotherapy. Exploratory factor analysis of three alliance measures. *Journal of Consulting and Clinical Psychology*, 64, 1326- 1336.
- Hedges, L.V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.
- Hedges, L.V., Tipton, E., & Johnson, M.C. (2010). Robust variance estimation in metaregression with dependent effect size estimates. *Research Synthesis Methods*. 1(1), 39-65.
- Higgins, J.P.T. & Thompson, S.G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistical Methods*, 21(11), 1539-58.
- Horvath, A. O. (1981). An exploratory study of the working alliance: Its measurement and relationship to outcome. Unpublished doctoral dissertation, University of British Columbia, Vancouver, Canada.
- Horvath, A. O., & Bedi, R. P. (2002). The alliance. In J. C. Norcross (Ed.), *Psychotherapy relationships that work* (pp. 37-69). New York: Oxford University Press.
- Horvath, A. O., Del Re, A., Fluckiger, C., & Symonds, D. B. (2011). Alliance in individual psychotherapy, *Psychotherapy*, 48(1), 9-16.
- Horvath, A. O., & Greenberg, L. (1986). The development of the Working Alliance Inventory: A research handbook. In L. Greenberg and W. Pinsoff (Eds.) *Psychotherapeutic processes: A research handbook*, New York: Guilford Press.
- Horvath, A. O. & Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy.

Journal of Consulting and Clinical Psychology, 61(4), 561-573.

Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology*, 38, 139–149.

Karver, M., Handelsman, J.B., Fields, S., & Bickman, L. (2006). Meta-analysis of therapeutic relationship variables in youth and family therapy: The evidence for different relationship variables in the child and adolescent treatment outcome literature. *Clinical Psychology Review*, 26(1), 50-65.

Kessler, R. C., Demler, O., Frank, R. G., Olfson, M., Pincus, H. A., Walters, E. E. & Zaslavsky, A. M. (2005). Prevalence and treatment of mental disorders, 1990 to 2003. *New England Journal of Medicine*, 352(24), 2515-2523.

Kessler, R. C., Angermeyer, M., Anthony, J. C. et al. (2007). Lifetime prevalence and age-of-onset distributions of mental disorders in the World Health Organization's World Mental Health Survey Initiative, *World Psychiatry*, 6, 168-176.

Littell, J.H., Campbell, M., Green, S.J., & Toews, B. (2007). Screening and data extraction forms: Systematic review and meta-analysis of effects of Multi-systemic Therapy (MST). Unpublished paper, Bryn Mawr, PA: Bryn Mawr College.

Luborsky, L., Crits-Christof, P., Alexander, L., Margolis, M., & Cohen, M. (1983) Two helping alliance methods for predicting outcomes of psychotherapy: A counting signs vs. a global rating method. *Journal of Nervous and Mental Disease*, 171, 480-491.

Marmar, C. R., Horowitz, M. J., Weiss, D. S., & Marziali, E. (1986). The development of the therapeutic alliance rating system. In L. S. Greenberg & W. M. Finsif (Eds.), *The psychotherapeutic process: A research handbook* (pp. 367-390). New York: Guilford. (1990, June). Reliability and validity of the three versions of the California Psychotherapy Alliance Scales (CALPAS). Paper presented at the meeting of the Society for Psychotherapy Research.

Martin, D. J., Garske, J. P., & Davis, K. M. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta analytic review. *Journal of Clinical and Consulting Psychology*, 68, 438–450.

McLeod, B.D. (2011). Relation of the alliance with outcomes in youth psychotherapy: A meta-analysis. *Clinical Psychology Review*, 31(4), 603-616.

Moher D., Liberati, A., Tetzlaff, J., Altman, D.G., & The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Med*, 6(7): e1000097. doi:10.1371/journal.pmed.1000097.

Newman, B. M., & Newman, P. R. (2011). *Development through life: A psychosocial*

approach. Wadsworth Publishing Company.

- Palmer, T., Peters, J., Sutton, A.J., & Moreno, S. (2008). Contour enhanced funnel plots for meta-analysis. *The Stata Journal*, 8; 242-254.
- Parker, K. (2012, March, 15). The Boomerang generation: Feeling ok about living with mom and dad, *PewResearchCenter*, retrieved from <http://www.pewsocialtrends.org/2012/03/15/the-boomerang-generation/>
- Pottick, K.J., Bilder, S., Vander Stoep, A., Warner, L.A., & Alvarez, M.F. (2008). US Patterns of Mental Health Service Utilization for Transition-Age Youth and Young Adults, *The Journal of Behavioral Health Services & Research*, 35(4), 373-389.
- Raue, P.J., Castonguay, L.G., & Goldfried, M. R. (1993). The working alliance: A comparison of two therapies, *Psychotherapy Research*, 3, 197-207.
- Rogers, C.R. (1951). *Client-centered therapy*. Boston, MA: Houghton-Mifflin.
- Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology*, 21(2), 95.
- Rosenzweig, S. (1936). Some implicit common factors in diverse methods of psychotherapy. *American Journal of Orthopsychiatry*, 6, 412-415.
- Schoeni, R. & Ross, K. (2004). Family support during the transition to adulthood. In R.A. Settersten, Jr, F.F. Furrsternberg, Jr., & R.G. Rumbaut, (Eds.), *On the frontier of adulthood: Theory, research, and public policy* (pp. 396-416). Chicago: University of Chicago Press.
- Shea, B.J., Grimshaw, J.M., Wells, G.A., Boers, M., Andersson, N., Hamel, C., Porter, A.C., Tugwell, P. Moher, D., & Bouter, L.M. (2007). Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews, *BMC Medical Research Methodology*, 7, 10.
- Shirk, S.R., & Karver, M.S. (2003). Prediction of treatment outcome from relationship variables in child and adolescent therapy: A meta-analysis. *Journal of Consulting and Clinical Psychology*, 71(3), 452-464.
- Shirk, S.R., Karver, M.S. & Brown, R. (2011). The alliance in child and adolescent psychotherapy, *Psychotherapy*, 48(1), 17-24.
- Suh, C., Strupp, H. & O'Malley, S. (1986). The Vanderbilt process measures: The Psychotherapy Process Scale (VPPS) and the Negative Indicators Scale (VNIS). In L. Greenberg and W. Pinsof (Eds.), *The psychotherapeutic process: A research handbook*, (pp. 285-323). New York: Guilford.

- The Jobless Young: Left Behind (2011). Retrieved January 16, 2013 from <http://www.economist.com/node/21528614>
- Tichenor, V., & Hill, C. E. (1989). A comparison of six measures of working alliance. *Psychotherapy, 26*, 195-199.
- U.S. Government Accountability Office (2008, June). Young adults with serious mental illness: Some states and federal agencies are taking steps to address their transition challenges. (Publication No. GAO-08-678). Retrieved from GAO Reports main page via GPO Access database <http://www.gao.gov/new.items/do8678.pdf>.
- U.S. Substance Abuse and Mental Health Services Administration (2012), *Results from the 2010 National Survey on Drug Use and Health: Mental Health Findings*, NSDUH Series H-42, HHS Publication No. (SMA) 11-4667. Rockville, MD: Substance Abuse and Mental Health Services Administration.
- Wallace, B. C., Trikalinos, T. A., Lau, J., Brodley, C., & Schmid, C. H. (2010). Semi-automated screening of biomedical citations for systematic reviews. *BMC Bioinformatics, 11*, 55.
- Wampold, B.E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Mahwah: Erlbaum.
- Wang, P. S., Angermeyer, M., Borges, G., Bruffaerts, R., Chiu, W. T., De Girolamo, G., & Üstun, T. B. (2007). Delay and failure in treatment seeking after first onset of mental disorders in the World Health Organization's World Mental Health Survey Initiative. *World Psychiatry, 6*(3), 177-185.
- Weinberger, J. (1995). Common factors aren't so common: The common factors dilemma. *Clinical Psychology: Science and Practice, 2*(1), 45-69.
- Zetzel, E.R. (1956). Current concepts of transference *International Journal of Psychoanalysis, 37*, 369-376.

6 Appendices

6.1 APPENDIX 1: SCREENING AND DATA EXTRACTION FORM

Version of 6/12/13:

Study ID:

Data	Authors or Report ID

Level 1: Initial Screening

1. Is this paper about the therapeutic alliance, working alliance, working relationship, or therapeutic relationship?

- Yes
- No [STOP]
- Can't tell

2. Is this paper about associations between the alliance and outcomes?

- Yes
- No [STOP]
- Can't tell [RETRIEVE FULL TEXT]

3. Is this paper about the alliance and outcomes for young adults aged 18-34?

- Yes
- No [STOP]
- Can't tell [RETRIEVE FULL TEXT]

4. What type of study design is this? (check all that apply)

- Observational (by therapist, client, and/or observer)
- Interview (quantitative)
- Longitudinal
- Interview (qualitative) (not eligible)
- Cross-sectional (not eligible)
- Self-report (not eligible)
- Descriptive or case study (not eligible)
- Can't tell

Level 2: Eligibility Decisions

1. Does this study include valid alliance measures?

- Yes
- No
- Can't tell

2. Does this study include valid outcome measures?

- Yes
- No
- Can't tell

3. Are results available specifically for client participants age 18-34?

- Yes (separate data provided for this age group)

- No, results are grouped in a larger age range or unspecified age range (e.g., “adults,” or “adolescents”)
- No
- Can’t tell

4. Does this study include outcome measures (e.g., depression, attendance/utilization, symptom improvement)?

- Yes
- No
- Can’t tell

5. Is the treatment in this study face-to-face psychotherapy performed by a licensed mental health practitioner or supervised participant in an accredited mental health training program and not excluded forms of counseling such as guidance counseling or mentoring?

- Yes
- No
- Can’t tell

Level 3: Data Extraction: Study Level

Research methods

1. Is sampling method -

- Probability
- Non-probability
- Other (specify)

2. What population was sampled for patients?

- From clinic or practice
- Other (specify)

3. If grouped, how are participants grouped (check all that apply)?

- Not grouped
- By age

- By therapist(s)
- By psychological condition
- By setting
- By technique/treatment type
- Can't tell
- Other (specify)

4. What population was sampled for psychotherapists?

- From hospitals, clinics, group practices, individual psychotherapy offices
- From existing survey data set
- Other (specify)

5. Who recruited/identified client participants?

- Research staff
- Provider staff
- Can't tell
- Other (specify)

6. Who recruited/identified psychotherapist participants?

- Research staff
- Provider staff
- Can't tell
- Other (specify)

7. Who recruited/identified observer participants?

- Research staff
- Provider staff
- Can't tell
- Other (specify)

8. How many age groups were included?

- 18-34 only
- 18-34 plus other adults
- 18-34 plus younger participants

- Other(s) (please specify)

Settings

9. Location of observations (check all that apply)

- Office setting (practitioner office)
- Home
- Research setting
- Can't tell (please specify)

10. Location details –observations/interviews took place in

- Office setting (practitioner office)
- Home
- Research setting
- Can't tell (please specify)

11. Location details – geographic

- City/region
- State/province
- Country (please specify)

Treatment Characteristics

12. Type of therapy

- Psychodynamic
- CBT
- Brief Treatment
- Substance abuse
- Other (specify)

13. Number/duration/hours of treatment/sessions

	Min	Max	Mean	SD	Pg# & Notes
Duration of services (number of sessions)					

Duration of services (duration of treatment in months)					
Duration of services (hours of treatment)					
Total number/duration/ hours of contact					

14. Sample size

14.1 Clients

Number of cases	Client	Total	Pg# & Notes
Clients referred to study			
Clients too ill			
Clients eligible			
Declined			
Data incomplete			
Clients - completed study			

14.2 Therapists

Number of cases	Total	Pg# & Notes
Therapists referred to study		

Therapists - completed study		
------------------------------	--	--

14.3 Psychotherapist/Client pairs (note in Psychotherapist/Client Order)

Number of cases	Psycho-therapist/Client	Total Psycho-therapists	Total Clients	Pg# & Notes
Referred to study				
Consented				
Completed study				

15. Sample Characteristics

15.1 Client demographic characteristics

		Total sample	Pg# & Notes
Gender	% Male		
Client age	Mean (sd)		
	Min and Max		
Race/ethnicity	% Caucasian		
	% Afr/Amer		
	% Hisp/Latino		
	% Asian/Pacific		
	% Other		

Socioeconomic status	% graduate level education		
	% college educ		
	% college student		
	% employed no college		
	% receive public aid		
Other sample characteristics			

15.2 Psychotherapist demographic characteristics

		Total sample	Pg# & Notes
Gender	% Male		
Therapist age	Mean (sd)		
	Min and Max		
Race/ethnicity	% Caucasian		
	% Afr/Amer		
	% Hisp/Latino		
	% Asian/Pacific		
	% Other		
Socioeconomic status	% Master's level		

		Total sample	Pg# & Notes
	% Master's level in training		
	% PhD		
	% PhD in training		
Therapist experience (years, etc.)	Mean (sd)		
	Min and Max		
Other sample characteristics			

16. Is there any information on clustering (info in text describing multiple people seeing same therapist, etc.)

Yes (describe)

No

Not sure

Level 4: Alliance and Outcome measures

1. When were data collected?

Alliance (A) or Outcome (O)	Observations – Pre-tx/During-tx / at what interval?	Client	Psycho-therapist	Observer

2. Who collected data?

- Psychological staff
- Research staff
- Other personnel (specify)
- Self-reports
- No interviews
- Can't tell

3. Were data collected in the same manner for all alliance measures?

- Yes
- No (what were the differences?)
- Can't tell

4. Were data collected in the same manner for all outcome measures?

- Yes
- No (what were the differences?)
- Can't tell

Alliance Measures

Instructions: Please enter measures in the order in which they are described in the report. Note that a single measure can be completed by multiple sources and at multiple points in time (data from specific sources and time-points will be entered later).

#	Topic	Reliability & Validity	Format	Direction	Source	Mode Admin	Pg# & notes
1	Code: Definition:	Info from: ○ Other samples This sample Unclear	○ Dichotomy Continuous	High score or event is ○ Positive ○ Negative ○ Can't tell	○ Therapist ○ Client ○ Observer ○ Unclear	○ Self-report ○ Researcher ○ Other ○ Don't know	

#	Topic	Reliability & Validity	Format	Direction	Source	Mode Admin	Blind?	Pg# & notes
	Alliance Measure							

#Code:	Info from:	○ Dichotomy	High score	○ Therapist	○ Self-report	○ Yes	
Definition:	○ Other samples	Continuous	or event is	○ Client	○ Researcher	○ No	
	This sample		○ Positive	○ Observer	○ Other	○ Can't tell	
	Unclear		○ Negative	○ Unclear	○ Don't know		
			○ Can't tell				

Outcome Measures

Instructions: Please enter measures in the order in which they are described in the report. Note that a single measure can be completed by multiple sources and at multiple points in time (data from specific sources and time-points will be entered later).

#Topic	Reliability & Validity	Format	Direction	Source	Mode Admin	Pg# & notes
Outcome Measure						
#Code:	Info from:	<input type="radio"/> Dichotomy <input type="radio"/> Continuous	High score or event is <input type="radio"/> Positive <input type="radio"/> Negative <input type="radio"/> Can't tell	<input type="radio"/> Therapist <input type="radio"/> Client <input type="radio"/> Observer <input type="radio"/> Unclear	<input type="radio"/> Self-report <input type="radio"/> Researcher <input type="radio"/> Other <input type="radio"/> Don't know	
Definition:	<input type="radio"/> Other samples <input type="radio"/> This sample <input type="radio"/> Unclear					

#Topic	Reliability & Validity	Format	Direction	Source	Mode Admin	Pg# & notes
Outcome Measure						

#Code:	Info from:	○ Dichotomy	High score	○ Therapist	○ Self-report	
	○ Other samples	Continuous	or event is	○ Client	○ Researcher	
Definition:	This sample		○ Positive	○ Observer	○ Other	
	Unclear		○ Negative	○ Unclear	○ Don't know	
			○ Can't tell			

Outcome data

Please enter outcome data in the tables provided below. Enter dichotomous outcomes first, then continuous outcomes. Outcome # refers to the measures described above.

Dichotomous outcome data

Outc #1	Days/weeks since start of treatment	Source	Valid Ns	Means	SDs	Statistics	Pg# & notes
		○ Client	Client				
		○ Therapist	Therapist				
		○ Observer	Observer				

Continuous outcome data

Outc #1	Days/weeks since start of treatment	Source	Valid Ns	Means	SDs	Statistics	Pg# & notes
	○	○ Client	Client				
		○ Therapist	Therapist				
		○ Observer	Observer				

Level 5: Study quality standards

1. Avoidance of **performance bias** (no differences confounded with alliance scores)

- Low Risk
- High Risk
- Unclear Risk

2. Avoidance of **attrition bias** (losses to follow up less than or equal to 20%)

- Low Risk
- High Risk
- Unclear Risk

3. Avoidance of **detection bias** (assessor unaware of alliance scores when collecting outcome measures)

- Low Risk
- High Risk
- Unclear Risk

4. **Standardized observation periods** (alliance and outcome measures collected from each case at a fixed point in time)

- Low Risk
- High Risk
- Unclear Risk

5. **Validated alliance measures** (use of instruments with demonstrated reliability and validity in this sample or similar samples OR use of public agency administrative data, behavioral, or biologic measures)

- Low Risk
- High Risk
- Unclear Risk

6. **Validated outcome measures** (use of instruments with demonstrated reliability and validity in this sample or similar samples OR use of public agency administrative data, behavioral, or biologic measures)

- Low Risk
- High Risk
- Unclear Risk

7. **Conflicts of interest** (e.g., researchers or data collectors would benefit if results favored specific alliance measures)

- Low Risk
- High Risk
- Unclear Risk

8. Comments:

7 Contribution of authors

SJG identified the need for a review on this population. SJG drafted the protocol, with substantive and methodological input from JHL. SJG and KH created the search strategy. SJG and ETS wrote the methods sections of the protocol. SJG developed the coding forms based on forms originally designed by Littell, Campbell, Green, Schaffner, and Towes (2007).

8 Declarations of interest

The authors have no vested interest in the outcomes of this review, nor any incentive to represent findings in a biased manner.

9 Sources of support

9.1 INTERNAL SOURCES

2012 Rivitz Award for Best Dissertation Proposal to Stacy J. Green from the Bryn Mawr College Graduate School of Social Work and Social Research

9.2 EXTERNAL SOURCES

Internal sources

- No sources of external support provided