



REFERENCE ONLY

UNIVERSITY OF LONDON THESIS

Degree PhD

Year 2005

Name of Author G. EMBRY A. J.

COPYRIGHT

This is a thesis accepted for a Higher Degree of the University of London. It is an unpublished typescript and the copyright is held by the author. All persons consulting the thesis must read and abide by the Copyright Declaration below.

COPYRIGHT DECLARATION

I recognise that the copyright of the above-described thesis rests with the author and that no quotation from it or information derived from it may be published without the prior written consent of the author.

LOANS

Theses may not be lent to individuals, but the Senate House Library may lend a copy to approved libraries within the United Kingdom, for consultation solely on the premises of those libraries. Application should be made to: Inter-Library Loans, Senate House Library, Senate House, Malet Street, London WC1E 7HU.

REPRODUCTION

University of London theses may not be reproduced without explicit written permission from the Senate House Library. Enquiries should be addressed to the Theses Section of the Library. Regulations concerning reproduction vary according to the date of acceptance of the thesis and are listed below as guidelines.

- A. Before 1962. Permission granted only upon the prior written consent of the author. (The Senate House Library will provide addresses where possible).
- B. 1962 - 1974. In many cases the author has agreed to permit copying upon completion of a Copyright Declaration.
- C. 1975 - 1988. Most theses may be copied upon completion of a Copyright Declaration.
- D. 1989 onwards. Most theses may be copied.

This thesis comes within category D.

This copy has been deposited in the Library of UCL

This copy has been deposited in the Senate House Library, Senate House, Malet Street, London WC1E 7HU.

Structural and symptomatic change in
psychoanalysis and psychodynamic psychotherapy:
A quantitative study of process, outcome, and attachment

Andrew James Gerber

Thesis submitted for the degree of Doctor of Philosophy
University College London
2004

UMI Number: U592082

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI U592082

Published by ProQuest LLC 2013. Copyright in the Dissertation held by the Author.
Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against
unauthorized copying under Title 17, United States Code.



ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

In memory of George Moran

Who inspired this work in every way

ABSTRACT

This thesis describes a quasi-experimental study exploring psychotherapeutic process and outcome in 25 young adults sequentially assigned to psychoanalysis (n=14) or psychodynamic psychotherapy (n=11) at the Anna Freud Centre in London, England. Analysts reported process using a novel 899-item questionnaire, the Young Adult Weekly Rating Scale (YAWRS). Patients were assessed by an independent psychiatrist at intake, termination, and at 18 month intervals after intake and termination with Main and Goldwyn's Adult Attachment Interview (AAI) and on a host of symptomatic and diagnostic measures.

The patients suffered from depression, anxiety, and personality disorders. Over the course of treatment (6 months to 8 years long), 12 of 19 patients (with adequate data) improved symptomatically on an aggregate measure. Ten of 12 improvers were in the psychoanalysis group, suggesting that it is a more effective treatment in this population. Data from 1,314 YAWRS questionnaires were factor analysed and used to test hypotheses from the psychotherapy process literature. In the first year of psychoanalysis (as compared with psychodynamic psychotherapy), higher scores on therapist dynamic technique, patient dynamic material, and negative patient transference were found. In the combined sample, higher scores in the first year on therapist dynamic technique, patient dynamic material, and discussion of contract were predictive of positive outcome.

The AAI classifies patients according to security of "state of mind with respect to attachment" from narratives about early life relationship experiences. Our results show a high proportion of secure classifications at initial assessment and, in successful treatments, a movement towards a preoccupied-entangled attachment pattern which began to resolve by termination. We propose that the AAI be used to measure both structural health and regression/transference neurosis, which must occur and then resolve for treatment to succeed. Further research using the YAWRS and AAI is proposed.

TABLE OF CONTENTS

Abstract.....	iii
Table Of Contents.....	iv
Appendices	vii
List Of Figures And Tables.....	viii
Acknowledgements	x
Chapter 1. Review of psychoanalytic outcome research.....	11
1.1 Introduction.....	11
1.2 Methodological and scientific challenges of psychoanalytic research	12
1.3 Historical review of psychoanalytic research	26
1.4 Current answers to major questions	34
1.5 Conclusion	51
Chapter 2. Psychotherapy research methodology.....	53
2.1 Introduction.....	53
2.2 Major themes within psychotherapy research	54
2.3 Theoretical issues in psychotherapy research	64
2.4 Practical solutions to methodological issues.....	90
2.5 Conclusion	114
Chapter 3. Description and assessment of Young Adult sample and treatment using standardised measures.....	115
3.1 Introduction.....	115
3.2 Methods.....	122
3.3 Results.....	145
3.4 Discussion.....	157
3.5 Conclusion	172

Chapter 4. The Young Adult Weekly Rating Scale: Measure development.....	175
4.1 Introduction.....	175
4.2 Methods.....	181
4.3 Results.....	193
4.4 Discussion.....	196
4.5 Conclusion.....	226
Chapter 5. The Young Adult Weekly Rating Scale: Psychoanalysis versus psychodynamic psychotherapy	228
5.1 Introduction.....	228
5.2 Hypotheses	239
5.3 Methods.....	240
5.4 Results.....	242
5.5 Discussion.....	247
5.6 Conclusion	258
Chapter 6. The Young Adult Weekly Rating Scale: Psychoanalytic and psychotherapy outcome	260
6.1 Introduction.....	260
6.2 Hypotheses	272
6.3 Methods.....	274
6.4 Results.....	276
6.5 Discussion.....	283
6.6 Conclusion	298
Chapter 7. Adult Attachment Interview.....	299
7.1 Introduction.....	299
7.2 Hypotheses	326
7.3 Methods.....	328

7.4	Results.....	331
7.5	Discussion.....	343
7.6	Conclusion	356
	References.....	441

APPENDICES

Appendix 3.1. Treatment parameters	358
Appendix 3.2: Raw assessment data.....	360
Appendix 3.3. Assessment measures	374
Appendix 3.4. Assessment data change scores	377
Appendix 3.5. Variables associated with overall improvement status.	381
Appendix 4.1. The Young Adult Weekly Rating Scale	383
Appendix 4.2a. Results of YAWRS subsection factor analyses.....	402
Appendix 4.2b. Results of subsection summary scale factor analysis.....	419
Appendix 4.3a. Formulae for calculation of factors from factor analyses (FAC).....	421
Appendix 4.3b. Formulae for calculation of subsection summary scores (SSC).	431
Appendix 4.3c. Formulae for calculation of global factors	437
Appendix 4.3d. Formulae for calculation of scales for hypothesis testing.	438
Appendix 7.1. AAI classification data	439

LIST OF FIGURES AND TABLES

Table 3.1. Patients in intensive treatment	123
Table 3.2. Patients in non-intensive treatment.....	124
Table 3.3. Clinical cutoff, standard error of differences, and sample information.	142
Table 3.4. Summary of attendance data for intensive and non-intensive subjects.	144
Table 3.5. Summary of assessment scales at initial, follow-along, termination, and follow-up for intensive and non-intensive patients.....	146
Table 3.6. Summary of initial-only assessment scales for intensive and non-intensive patients.	147
Table 3.7. Numbers of intensive and non-intensive subjects in the clinical range at initial, termination, and follow-up.....	147
Table 3.8. Numbers of intensive and non-intensive subjects in the clinical range at initial-only.	147
Table 3.9. Numbers of intensive and non-intensive subjects with DSM-III-R Axis I and Axis II diagnoses at initial, termination, and follow-up.	148
Figure 3.1a-d. Mean scale values at initial assessment and termination.....	151
Figure 3.2a-d. Mean scale values at initial assessment and termination by treatment intensity	152
Figure 3.3a-d. Mean scale values at initial assessment, follow-along, and termination. ..	153
Figure 3.4a-d. Mean scale values at initial assessment, termination, and follow-up	154
Table 3.10. Numbers of subjects with change patterns on initial versus termination assessments.....	155
Table 4.1a-b. Numbers of YAWRS forms collected by quarter year from beginning of treatment.	193
Table 4.2. Multiple factor solutions from subsection factor analyses.....	195
Table 4.3. Global factors from subsection summary factor analysis.	196

Table 5.3. Composition of therapist-based hypothesis scales.....	243
Table 5.4. Composition of patient-based hypothesis scales.....	245
Table 5.5. T-tests: Psychoanalysis versus psychodynamic psychotherapy	246
Table 5.6. HLM analysis: Psychoanalysis versus psychodynamic psychotherapy	247
Table 6.3. Composition of therapist-based hypothesis scales.....	277
Table 6.4. Composition of patient-based hypothesis scales.....	278
Table 6.5. Composition of interaction-based hypothesis scales.	279
Table 6.8. T-tests: Improvers versus non-improvers	281
Table 6.9. HLM analysis: Improvers versus non-improvers	282
Table 7.1a-b. AAI classifications.	333
Table 7.2a-b. AAI RF and probable experience subscales at intake and termination.	334
Table 7.3a-b. AAI probable experience subscales at intake and termination.	335
Table 7.4. AAI subscale initial means and change means.....	336
Figure 7.1. Positive SOM scale in follow-along analysis.....	340
Figure 7.2. Positive SOM scale in follow-up analysis.....	341
Table 7.5. YAWRS factors by Initial AAI classification.	342
Table 7.6. Correlations between YAWRS factors and Initial AAI subscales.	342
Table 7.7. YAWRS factors by AAI change classification.	343
Table 7.8. Correlations between YAWRS factors and change in AAI subscales.....	343

ACKNOWLEDGEMENTS

During the 10 year gestation of this thesis, many people have provided invaluable assistance. The project could not have been completed without tireless help from Clare Welch, Julia Curl, Liz Allison, and Ros Bidmead. Financial support was provided by the Anna Freud Centre, the Granville family, University College London, and the International Psychoanalytical Association. I would especially like to thank Anna Higgitt for correcting a draft of this thesis, and, along with Anthony Bateman, devoting her time and expertise to the study. I am also grateful to the many AAI coders and research colleagues who helped along the way, particularly John Crosse and Saul Hillman.

This project grew out of a forward-thinking and dedicated group of psychoanalysts who came together to show that psychoanalysis can be the subject of rational empirical research. Under the leadership of Ann Marie Sandler, this team (Marion Burgner, Luigi Caparrotta, Rosemary Davies, Rose Edcumbe, Julia Fabricius, Ann Harrison, Brian Martindale, Duncan McLean, Rosine Perelberg, Joan Schachter, Maria Tallandini, Sally Weintrobe, and Anne Zachary) welcomed and educated me by openly sharing their work. I am indebted as well to their patients and hope that the intrusion they accepted in their treatments is made worthwhile by a contribution that will help future young adults.

I am fortunate and honoured to have been mentored by an extraordinary set of psychoanalysts and researchers. Joseph Sandler, Mary Target, and Howard and Miriam Steele inspired me with their examples. My greatest debt is to Professor Peter Fonagy who guided this research at every stage. Through his intellect, enthusiasm, and generosity, he has shaped the way I view the world and inspired the work I have chosen.

Finally, I am grateful to my parents for supporting me in my efforts and to my wife, Andrea, for her love and patience. She calmly listened to my struggles, supplied me with an inexhaustible list of fresh ideas, and brought me the confidence and joy I needed to bring this task to its completion.

CHAPTER 1. REVIEW OF PSYCHOANALYTIC OUTCOME RESEARCH

1.1 Introduction

It is a long-standing and strangely well-accepted myth that the therapeutic practice of psychoanalysis* has not been the subject of extensive empirical research. In the last decade, at least ten thorough reviews of psychoanalytic research have been published, describing dozens of large scale efforts over the past century to systematically measure the effects of psychoanalysis (Bachrach, Galatzer-Levy, Skolnikoff, & Waldron, 1991; Doidge, 1997; Fisher & Greenberg, 1996; Fonagy, Kächele, Krause, Jones, & Perron, 1999; Fonagy et al., 2001; Galatzer-Levy, Bachrach, Skolnikoff, & Waldron, 2000; Kantrowitz, 1997; Leuzinger-Bohleber & Target, 2002; Roth & Fonagy, 1996; in press). As each of these reviews amply documents and will be summarized later in this chapter, there are many substantial obstacles to performing scientifically rigorous research on psychoanalysis. These obstacles, including the length and heterogeneity of the therapy being studied, as well as the difficulty of measuring pertinent process and outcome variables, are typically blamed for flaws within the studies conducted, and the relative lack of consequence that the research has had in influencing clinical practice (Fonagy et al., 2001; Sandell et al., 2000). However, understanding the specifics of the research that has been done and the limitations that have kept it from being more widely known, is essential to a contemporary approach to psychoanalytic research.

* In this thesis, “psychoanalysis” will be used to refer to the treatment as recognized by the International Psychoanalytical Association (IPA), performed by an IPA-accredited psychoanalyst typically four to five times per week with the patient lying on a couch. “Psychodynamic psychotherapy” will denote psychotherapy that follows similar theoretical principles but is conducted less frequently. See Chapter 2 for further theoretical and clinical discussion of the differences between these treatments.

1.2 Methodological and scientific challenges of psychoanalytic research

1.2.1 *Epistemology of psychoanalytic research*

An important starting point for psychoanalytic research projects, that is even more central today than it was when such studies began, is to ask why psychoanalysis should be studied scientifically at all. Historically, researchers felt that applying the objectivity, methodological rigor, and reproducibility of science to psychoanalysis would lead to progress in our theoretical understanding of the treatment and our ability to recommend and provide more effective care to patients. Observing the benefit which scientific investigation has had in medicine and technology, it was reasonable to hope that the effect on the practice of psychoanalysis would be similar (Galatzer-Levy et al., 2000). Recently, with increasing focus on limited resources in health care and the need for cost-effective approaches to treatment, researchers have discussed the benefits of psychoanalytic and psychotherapy research in these terms as well (Fonagy, 1999a; Gabbard, Lazar, Hornberger, & Spiegel, 1997; Lazar, 1997; Richardson, 2001).

It has never been universally agreed upon, though, that the scientific approach is appropriate to the study of psychoanalysis. The French psychoanalytic community, to this day, has expressed great reservations about relying on scientific methods to describe a process that many believe to have a distinct epistemology. Summarizing this position, Perron expresses the fear that attempts at quantification of analytic material would lead to “fragmentation of the materials, no subsequent statistical calculation, however sophisticated, being able to restore the lost unity” (1999, p. 16). He also notes a widely held concern that audio or video recording of analytic sessions gravely disturbs the transference-countertransference relationship, and thus changes the very process it seeks to measure.

Fonagy, a major advocate and practitioner of psychoanalytic research, outlines ways in which the analytic community has made research difficult by not complying with

certain requirements of the scientific perspective. He sees the greatest problem as the confusing absence of shared theoretical and clinical assumptions among analysts and the lack of a common and well-defined language. Fonagy believes this problem has only worsened in recent years as theories fragment and theory and clinical practice develop independently with only tenuous links to one another. As evidence, he points to the fact that the rates with which analysts cite one another in peer-reviewed journals is steadily declining (1999d).

By any standards, the general problems faced by psychoanalytic research are daunting.

The systematic study of psychoanalysis is intrinsically difficult. Psychoanalysis is a complex, subtle, difficult-to-observe activity that extends over many years; the assessment of its effects may take decades. Second, those who choose to become psychoanalysts are generally more interested in understanding individual psychology in depth, not in quantitative methodologies. Third, the long history of psychoanalytic claims to scientific status has interfered with systematic exploration. In order to undertake such exploration a first painful step is to recognize that many of these claims are ill founded. Finally, the initial criticisms of psychoanalysis based on statistical analysis of studies of efficacy (Eysenck, 1952) were so clearly hostile and unfair that many analysts equate quantitative studies with negative attitudes toward psychoanalysis, and many continue to do so (Galatzer-Levy et al., 2000, p. ix).

Fortunately, as the studies described below illustrate, there is no shortage of investigators who have been determined to overcome these obstacles. We believe that by studying the successes and failures of these studies and by asking specifically what needs to be measured in order to construct a useful study, progress will continue to be made in the field.

1.2.2 Technical problems in psychoanalytic research

One of the questions that has emerged from 87 years of psychoanalytic research and commentary (the first outcome study was published by Isadore Coriat in 1917 under the title “Some statistical results of the psychoanalytic treatment of the psychoneuroses”) is how best to state the fundamental goals of this work. In 1943 Obendorf surveyed a

group of psychoanalysts with eight questions: (1) What percentage of your patients do you treat with psychoanalysis? (2) What percentage of these cases have prematurely terminated? (3) What percentage of treated cases have avoided hospitalization? (4) Should the analysis of patients who do not improve be terminated? (5) Do you believe in tapering off treatment? (6) What are your criteria for termination? (7) Do you distinguish analytic from symptomatic success? and (8) With which types of cases have you achieved the best result? Galatzer-Levy and colleagues summarize these questions in an elegant statement of what all researchers want to know, “What type of treatment is best suited to what kind of patient, suffering from what kind of illness, at what point in life, when treated by what kind of analyst, in what manner?” (2000, p. 53). Strupp, Schacht, and Henry formulated a similar concept in more technical language, “The principle of P-T-O congruence proposes that the intelligibility of psychotherapy research is a function of the similarity, isomorphism, or congruence among how we conceptualize and measure the clinical problem (P), the processes of therapeutic change (T), and the clinical outcome (O)” (1988, p. 7).

In accord with these statements of the goals of research, the technical problems faced by actual studies can be framed in terms of four questions: (1) How do we reliably and validly characterize patients according to diagnosis?, (2) How do we reliably and validly define the practice of psychoanalysis?, (3) How do we reliably and validly measure improvement in response to psychoanalysis over time?, and (4) How do we attribute change during treatment to the specific elements of psychoanalysis that we believe to be essential to change? Implicit in each of these questions is the task of coping with the enormous heterogeneity found in patient populations, methods of psychoanalytic treatment (including duration and frequency of therapy), ways in which patients change over time, and life experiences, both in and outside of the psychoanalysis, which might have been related to change. These questions not only summarize the challenges faced by

psychoanalytic research, but also provide a set of criteria on which research studies can be evaluated.

1.2.3 Initial evaluation of patients

An important part of virtually every psychoanalytic outcome study, has been to collect pre-treatment data on patients regarding demographics, life-circumstances, and psychopathology so as to better understand who is treated in psychoanalysis and whether these features can be used to predict treatment patterns or outcome. From the early years of the field, analysts have debated the scope of psychoanalysis, and the qualities that make an individual “analyzable” (Erle & Goldberg, 1979). This has long seemed to be an empirically answerable questions, requiring only that we know how to adequately categorize patients and measure which of them improves.

The nature of the data collected has closely mirrored the diagnostic frameworks in favor at the time and the type of information thought to be relevant to psychoanalytic treatment. For example, surveys during the early part of the century classified patients as “neurotic,” “psychotic,” or sometime “psychosomatic” (Fenichel, 1930; Jones, 1936; Knight, 1941). Demographic and life-experience data collected were often sparse, as psychoanalysis was believed by some to be effective independently of and despite the diversity in these variables. In the 1950’s, psychoanalytic researchers outgrew established diagnostic systems and measures of psychopathology, and became the leaders in developing such classifications. The Menninger Clinic, partly under the demands of their Psychotherapy Research Project (PRP), established a wide ranging set of evaluation scales still in use today, including psychological tests standardised by Rapaport and colleagues and Luborsky’s Health Sickness Rating Scale (HSRS; Wallerstein, 1986). Data about the patient’s life experience were collected in detailed narrative accounts, though systematic and quantitative measures lagged behind.

Yet another difficulty with attempts to use initial assessments of psychopathology to predict outcome is that no matter how careful the effort some aspects of pathology may not surface until after the treatment has begun. Despite extensive protocols for evaluation in the Menninger, New York, and Boston studies of psychoanalysis, retrospective diagnoses at termination appeared worse than had been suspected initially in as many as half the patients (Erle, 1979; Kantrowitz, 1993; Wallerstein, 1986). Though improved methods may reduce this problem, it is likely that at least some aspects of pathology cannot accurately be detected until psychotherapy is under way.

Researchers have been driven to develop new measures, both by the desire to describe patients more completely, and by the frustration of studies failing to predict outcome on the basis of initial diagnosis. Although they are routinely collected, few analytic researchers consider DSM-IV diagnoses sufficient for a psychoanalytic formulation, and have supplemented with their own systems including measurements of “ego strength” (1985a; Weber, Solomon, & Bachrach, 1985), personality (Fisher & Greenberg, 1977), quality of object relations and affect tolerance (Kantrowitz, 1993), “maladaptive interpersonal core patterns,” and attachment classification (Fonagy et al., 2001). In contemporary research, investigators usually aim to collect diagnostic information from multiple informants (e.g., patient, analyst, family), through multiple techniques (self-report, diagnostic interview, psychological testing, direct observation), in multiple domains (e.g., symptomatology, personality, character structure, affect regulation), and affecting different domains of functioning (e.g., work, social, family). Years of research in psychological assessment, particularly of children and adolescents, have shown that each perspective provides unique information and making multiple observations is the only solution to bias inherent in any one approach (Achenbach & McConaughy, 1997; Greenbaum, Dedrick, Prange, & Friedman, 1994; Kashani,

Orvaschel, Burk, & Reid, 1985). Many standard measures have been refined over the years and have well demonstrated reliability and validity.

Despite progress in the area of diagnostic assessment, many questions remain that have not been systematically addressed. To start, clinicians intuitively recognize that the point at which they have enough information to make a reasonable prediction as to the analyzability of a patient varies between patients, yet the timing of diagnostic assessments has not been systematically studied and is usually determined by practical constraints. Possible time points include the first few minutes of meeting the patient, the conclusion of a formal assessment period, the time at which a transference has begun to emerge, and the time at which transference is firmly established. Some analysts have proposed that a trial of analysis is the only accurate method for determining a patient's analyzability, and a study of the length of such a trial period could be informative (Galatzer-Levy et al., 2000).

The list of a patient's traits and capacities that might contribute to our ability to predict psychoanalytic outcome is long, and is limited only by the possibility of defining and operationalising such items. Galatzer-Levy (2000) suggests a partial list of traits that have not yet been adequately studied: transference readiness, capacity to modulate intensity of transference, capacity for free association, rigidity of defences, capacity to use interpretation, psychological mindedness, richness of association, and capacity for controlled regression. An ongoing challenge of methodologists will be to find ways to clearly define and measure these capacities such that researchers from multiple backgrounds are willing to apply them to diverse groups of subjects.

While analysts have traditionally focused on measures of psychopathology and mental structure as providing the clues for predicting outcome, they have been slower to develop categorizations and rigorous methods for measuring life experiences. Only in recent years are structured interviews for assessing childhood trauma becoming available

(Bifulco, Brown, & Harris, 1994) and these have not yet been applied to large psychoanalytic research projects. Demographic information, such as gender and socioeconomic status, are rarely used to stratify samples to account for variability in outcome.

A final challenge in the initial evaluation of a patient emerges from the suggestion that analysis must be viewed as a two-person process, depending as much on the characteristics of the patient as on the interdigitation with the analyst (Kantrowitz, 1993). Although, several studies have observed possible differences between outcome from candidate and graduate analysts (Erle, 1979), Sandell's study is the first major effort to look at a broad range of analyst characteristics in relation to process and outcome (2000). Relevant data about the analyst might include theoretical orientation and level of training. Far more difficult to measure, but possibly more important, are the characteristics of the patient-analyst match, such as similarity, complementarity, and the analyst's ability to invite transferences that need to be analysed (Galatzer-Levy et al., 2000).

1.2.4 Evaluation of analytic process

If a study purports to study the outcome of psychoanalysis, investigators must at some level decide how they define psychoanalysis and what means they have to determine whether this treatment took place. Galatzer-Levy suggests a set of ways to identify psychoanalysis that roughly corresponds to a developmental scale for conceptualizing and measuring the treatment: (1) the activity of psychoanalysts, (2) the practice derived from psychoanalytic theory, (3) an approach to patients and method of collecting patient data, (4) a set of techniques, and (5) an activity defined by a dyadic process (2000). In the early years of outcome research, the need to measure or even identify analytic process had not yet been established. This likely grew out of the insecurity of the analytic field in defining, either for themselves or for the scientific community, exactly what they did that was essential for patient care, and the fear of

researchers that if they sought to study this more carefully, they would alienate the clinical community whose cooperation they required. This began to change slowly in the 1950's with the advent of survey studies comparing various types of therapy that asked basic questions about session frequency, type of therapy performed, and length of treatment. A large survey conducted by the American Psychoanalytic Association was the first to query analysts about circumstances of termination (Hamburg et al., 1967).

The measurement of analytic process began with the Menninger PRP, which asked raters to review analyst process notes and treatment summaries and classify the type of therapy that was being performed (Wallerstein, 1986). The reliability and validity of these assessments was poor, though, and even within the study researchers used different classification, Wallerstein dividing interventions into "supportive" and "interpretive" and Kernberg classifying psychotherapy along the continuum of "supportive" to "expressive" (Kernberg, 1973). Confusion over these classifications likely contributed to the contradictory findings of separate analyses (L. Coyne, personal communication).

Surprisingly little effort was expended by researchers in studying what they meant by "psychoanalysis" until the past 20 years. Sandell (2000) sought to address the question by surveying analysts about their beliefs, and classifying the therapy they performed according to their responses to a questionnaire. Audio and videotaping of psychoanalytic sessions did not begin until the 1980's in Germany, and outside of that country is still uncommon and frowned upon by many in the analytic community. With the availability of analytic transcripts a set of process measures evolved and are described in more detail in Chapter 2.

A major trend in non-psychodynamic psychotherapy research in recent years has been the establishment of treatment manuals that specify the rules according to which a given psychotherapy is performed. The generally accepted assumption has been that only with such manuals are researchers able to clearly define treatment groups and conduct

experiments that compare such treatments. With audio or video recording of a representative sample of sessions, the application of adherence measures, derived from the manuals, is possible, further reassuring investigators that they are comparing the appropriate techniques. Many have argued that because psychoanalysis has resisted manualization, it has been left behind in the progress towards better outcome research and will not be able to reach the verifiability of an “empirically validated therapy.”

There is recent evidence, however, that even manualized therapies are practised differently than would be expected. Ablon and Jones (1998) measured process from psychotherapy transcripts of both cognitive and psychodynamic short-term therapies for depression using the Psychotherapy Process Q-sort (PQS, Jones & Pulos, 1993) and found that therapies of each type used techniques identified with the other approach. Furthermore, in both forms of treatment, positive outcome was associated with the extent to which the treatment matched the empirical prototype of psychodynamic psychotherapy (as defined by a wide range of process items, described in Chapter 4). A careful review of the data on so-called empirically validated therapies is not convincing in its exclusion of psychoanalysis, and process measures may be the route to its re-entry into the world of well validated research (Westen & Morrison, 2001).

Psychoanalytic outcome research requires several things of a good process measure. It should provide an operational definition of psychoanalysis, distinguishing it from psychodynamic psychotherapy and other psychotherapeutic modalities. A process measure should provide insight into the phases of analysis, with the possibility of relating these phases to periods of change and eventual outcome. Researchers have suggested broad categories for active elements of the analytic process, such as therapeutic contact, intervention, bond self-relatedness, and therapeutic realizations (Kolden & Howard, 1992) and a process measure should help validate these theories.

Due to the diversity of analytic process and the potential for effect on outcome, several other variables should be collected in systematic research. Because of the availability of trainees for research, many studies have used candidate or novice analysts as their treaters (Bachrach, 1993; Kantrowitz, 1993). Yet, other than Sandell's findings (2000), there is no reliable data on the effect of training on analytic outcome, calling into question the generalisability of many of the earlier studies. Similarly, there are no large studies of the effect of small changes in frequency (between three and five times a week), amount that the patient is asked to pay, policies for handling absences, vacations, and terminations, position of the patient (lying down vs. sitting up), just to name a few. As studies of process and outcome become more sophisticated, the enormous variety of techniques part of the analytic world will be available for scrutiny.

1.2.5 Evaluation of analytic outcome

Of all the challenges faced by psychoanalytic outcome research, none is greater than the task of defining a set of outcome measures that meaningfully captures change in response to psychoanalysis. The problem has many elements. First, even within the analytic community it has always been far from clear as to what the goals of psychoanalysis truly are. Dozens of books and hundreds of articles have debated this question and traced changes in the field's conceptualization of goals through different times and theoretical frameworks (Sandler & Dreher, 1996). Galatzer-Levy (2000) lists seven domains that might be considered important signs of change: (1) Symptom relief, (2) Development of insight, (3) Structural change, (4) Maturation of defensive operations, (5) Promotion of growth, development, values, (6) Improvement in quality of object relations, and (7) Occurrence of certain processes, such as catharsis, analysis of defence, analysis of transference, transformation of internal objects, and experience of appropriate parenting.

Even if it is accepted that there is an inherent diversity in the ways patients benefit from psychoanalysis, theorists disagree as to how change should best be measured. Perron, representing the French-speaking position on psychoanalytic research, asserts that “measurement of symptom reduction will not be enough. We know that symptoms are erratic, that if one disappears it may give way to another, that some symptoms are useful because they are part of defences and their careless destruction might be dangerous, etc.” (Perron, 1999, p. 18). In direct opposition, Sandell argues that if we believe that symptoms are “substitute formations” they would be activated by failure of treatment to change underlying structures, and are therefore a good proxy for measuring structural change (2000). The theory of structural change and proposed methods for measuring it, will be further discussed in Chapter 2.

Inextricably linked to the question of *what* should be captured in measurements of psychoanalytic outcome, are the questions of *how* and *when* to make these assessments, and *whom* to use as an informant. The findings from the assessment literature suggest that, just as for initial evaluation, there is no substitute for collecting data using as many techniques (e.g., self-report, interview, psychological testing, direct observation), from as many informants (e.g., patient, analyst, family), in as many domains (e.g., symptoms, personality, character structure, affect regulation), and in as many different areas of functioning (e.g., work, social, family), as possible. Each perspective provides unique information that is potentially valuable to the measurement of change.

Historically, studies of analytic outcome have very gradually come to the terms with the difficulties of measuring change and expanded the sophistication of their techniques. Until the 1950’s, research relied entirely on the analyst’s report of outcome, and investigated only a single, generic domain of “improvement.” The large scale survey of the American Psychoanalytic in 1952 continued to use only analysts’ reports, but made significant progress in differentiating between “symptom cure” (relatively rare,

characterizing only 25% of the sample), improved functioning (97% of the sample), and change in “character structure” (predicted by initial diagnosis) (Hamburg et al., 1967).

The greatest leap forward in psychoanalytic change methodology came about as a result of the Menninger PRP. Wallerstein and colleagues introduced several new techniques for reliably and objectively assessing change, based either on the application of psychological testing to subjects or the systematic coding of clinical evaluations and process notes from the patient’s chart. In both cases, ratings were assigned by researchers blind to clinical and treatment status of the patient, and thus unbiased in a way that primary clinicians could never be (Wallerstein, 1986). Scales used to rate psychological tests and patient charts focused on many of the theoretical constructs that are valued to this day: nature of conflict, ego factors and defence, capacities, motivation, and relationships. The Menninger study also introduced a unique method called “paired comparisons,” whereby instead of asking an investigator to assign an arbitrary number to a scale, he or she was asked to make thousands of individual choices as to whether the scale was more prevalent in one subject or another. Although, because of its time consuming nature, the method of “paired comparisons” has rarely been applied again, it is often hailed as one of the major contributions of the PRP (Kantrowitz, 1993).

Beginning with the Menninger project, many studies moved away from the traditional analyst-dependent and one-dimensional view of change. Studies at the Boston Psychoanalytic Institute also made use of psychological testing and rated several areas of change including symptoms, life functioning, relationships, and conflict (Knapp, Levin, McCarter, Wermer, & Zetzel, 1960; Sashin, Eldred, & Van Amerongen, 1975). The Columbia records project classified psychoanalytic change via the circumstances of termination, independent ratings of overall improvement, and independently assigned specific change scores (Bachrach, 1993). One of the most ambitious data collecting efforts in the assessment of change was led by Pfeffer in the development of a technique

for interviewing subjects years after termination to assess outcome (Oremland, Blacker, & Norman, 1975).

As the focus of psychoanalytic outcome research has shifted to sites in Europe over the past 20 years, investigators have continued to pursue objective measures from multiple perspectives, with occasional development of new standardised measures. A standard battery usually includes a set of self-report questionnaires measuring general psychiatric symptomatology, depression, anxiety, social adjustment, and life functioning. Sandell and colleagues in Sweden introduced Change after Psychotherapy Scales (CHAP) which combine measures of symptoms and adaptive capacity with a system for quantifying “self-insight” and “basic conflicts” (2000). Several studies in Western Europe, possessing good access to national health care utilization and work absenteeism records, have introduced these as objective and highly practical measures of analytic improvement.

Cierpka and Schneider in Germany founded a task force in 1992 to create a set of Operationalised Psychodynamic Diagnostics that is now in its third German edition and has been translated into several other languages, including English (OPD Task Force, 2001). This comprehensive and carefully manualized system includes four psychodynamic axes and one descriptive axis: (i) experience of illness and prerequisites for treatment, (ii) interpersonal relations, (iii) conflict, (iv) structure, and (v) mental and psychosomatic disorders. The Heidelberg-Berlin study currently under way makes use of the OPD for both initial assessment and for quantifying change (Grande, Rudolf, Oberbracht, Jakobsen, & Keller, 2004). Other measures that have begun to be applied in German outcome studies include Wallerstein’s Scales of Psychological Capacities (Huber & Klug, 2004; 1988) and Stiles’s Assimilation of Problematic Experiences (1992). Several more such instruments are currently in use or in development as part of psychotherapy process and outcome studies.

The proliferation of psychoanalytic outcome measures has done little to solve some of the fundamental questions still posed by researchers and critics. While it is generally accepted by investigators that independent ratings of patient outcome are essential for objectivity, clinicians justifiably point out that it is hard to conceive of a scale, or set of scales, that could take into consideration the complexity of change that every therapist follows on a day to day basis. Given this challenge, it is reassuring that research on therapist assessment has found it to be reliable and to correspond well to independent assessed outcome (Berzins, Bednar, & Severy, 1975; Mintz, Luborsky, & Christoph, 1979).

Starting with Eysenck (1952) skeptics of psychoanalytic outcome research have pointed out that spontaneous remission may account for a large part of change, and that without appropriate comparison groups, measurements of change are difficult to interpret. Because of the ethical difficulties of maintaining a non-treated control group and the practical difficulties of following a waiting list sample, it is rare for studies to have adequate comparison groups against which to judge the significance of change (Fonagy, 2001b). Perron points to the frequent case of a patient who concludes treatment with no noticeable improvement, but for whom we are justified to think that the situation may have been far worse without treatment (1999).

The ever increasing number of outcome measures exacerbates the problem, as it is difficult to interpret the comparability of change on any given measure (Fonagy et al., 2001). Researchers have been slow to adopt one another's scales, as each one conceptualizes change with a slightly different theoretical model. Particularly concerning is the possibility that measures of change may specifically tap domains that are close to those targeted by a particular therapy, giving falsely elevated impressions of change (Fonagy, 2001b). Only as a standardised battery emerges that is satisfactory to all

investigators and data are collected on a suitable range of patients in a range of therapies will these questions be addressed.

1.2.6 Linking change to analytic process

Even if outcome research is able to demonstrate that psychoanalysis is taking place and that patients are improving more than would be expected by chance, it remains to be shown that those elements we identify as specific to psychoanalysis are responsible for that change. Only very recently have studies begun to show that psychoanalysis is associated with longer lasting and more structural change than competing therapeutic modalities (Sandell et al., 2000). However, it is still not clear whether these differences are due to specific factors such as interpretive interventions, analysis of transference, and depth of analyst-patient interaction, or to non-specific factors, such as length of time spent with patient and level of training of the therapist. The only two techniques that offer to shed light on this question are the use of more specific control groups and the application of sophisticated measures of process. Control groups are limited by the difficulty in matching patients on all the variables that we consider pertinent and the impossibility of blinding patients and therapists, both of whom will have biases, to the nature of the therapy being done. Anna Freud quoted her father as having joked that the only truly satisfactory controlled study would be to treat the same person twice – once with analysis and once without (Pfeffer, 1959). Process measures will require years of development to find one that is broad enough to capture all the pertinent variables yet sufficiently sensitive to find what is related to outcome. However, such a measure holds the only promise of making causal links between process and outcome possible.

1.3 Historical review of psychoanalytic research

A complete historical review of psychoanalytic outcome research has not yet been written and is beyond the scope of this chapter. Psychoanalytic outcome studies can

broadly be divided into five time periods, each with its own focus: (1) early efforts with survey data (1917-1939), (2) the golden age of large scale American outcome research (1950-1959), (3) broader application of research with more specific research questions (1960-1979), (4) a shift to non-American studies of psychoanalysis (1980-2000), and (5) current research efforts (1990-present).

In the earliest period, Isador Coriat of the American Psychoanalytic Association (1917), Otto Fenichel of the Berlin Psychoanalytic Institute (1930), and Ernest Jones of the London Psychoanalytic Institute (1936), reported retrospective data on 93, 721, and 74 patients, respectively. These samples were composed of 70-80% neurotics with the remainder classified as psychotic or psychosomatic. Rates of “symptom cure” were approximately 50% in the neurotic group and significantly lower in the psychotic group. Similar results were reported in surveys by Franz Alexander at Chicago Psychoanalytic Clinic (1937), Kessel and Hyman at the New York Psychoanalytic (1933), and Robert Knight at the Menninger Clinic (1941).

The 1950’s witnessed at least 11 studies of psychoanalytic outcome, several of which were quite large and included a range of new methodologies. Studies by Heine (1953), Schjelderup (1955), Ellis (1957), and Orgel (1958) began to focus on the specific question of how psychoanalysis compares in effectiveness with alternative therapies or no therapy at all (Fisher & Greenberg, 1977). The most ambitious studies, however, were initiated during this time by the Central Fact-Gathering Committee of the American Psychoanalytic Association (Hamburg et al., 1967), the Menninger Clinic (Kernberg et al., 1972; Wallerstein, 1986), the Boston Psychoanalytic Institute (Knapp et al., 1960; Sashin et al., 1975), and the Columbia Psychoanalytic Center (Bachrach, 1993; Bachrach, Weber, & Solomon, 1985; 1985a; 1985b; Weber, Solomon et al., 1985).

Most notably, the Menninger Psychotherapy Research Project (PRP), collected a range of data, still unmatched to this day, on 42 patients in psychoanalysis or

psychodynamic psychotherapy and led to dozens of scientific papers. Of the 22 patients in psychoanalysis eight had very good outcome, five had reasonable but limited outcome, three had equivocal outcomes, and six were clear-cut failures, as judged by a host of objective measures. Of the 20 patients in psychotherapy, nine had very good outcome, three had reasonable, three had equivocal, and five were failures. Robert Wallerstein, the longterm director of the study, attributed the low improvement rate and difficulty in predicting improvement to the fact that despite using all available techniques, initial diagnosis was inadequate and in more than half of the psychoanalytic patients, major pathology was not initially revealed that would have been a contraindication to treatment. He also came to believe that contrary to prevailing theory, interpretation was not associated with “structural” change in these patients, more so than supportive techniques. In fact, the psychotherapy group, in which supportive interventions were found to be more common (though, supportive interventions were found to be a more significant component of psychoanalysis than had been expected), did better than expected and were not distinguishable from the psychoanalysis group (Wallerstein, 1986).

Kernberg and colleagues (1972) compiled the quantitative and statistical findings of the Menninger PRP, focusing on measures of “ego strength” and their relationship to outcome in the various patient groups. Ego strength was defined via factor analysis as (1) degree of integration, stability, and flexibility of intrapsychic structures (patterning of defences and anxiety tolerance and, implicitly, impulse control, thought organization, sublimatory capacity), (2) degree to which relationships are adaptive, deep, and gratifying of normal instinctual needs (quality of interpersonal relationships), and (3) degree to which malfunctioning of intrapsychic structures is manifested by symptoms (severity of symptoms). They reported that a high level of initial ego strength was a favorable prognostic sign, irrespective of the mode of treatment. Patients with low initial ego strength, who mostly fit into the diagnostic category of borderline personality disorder,

appeared to do best with “high-skill” therapists who focused on transference in treatment but also supplied supportive elements. Low initial quality of interpersonal relationships was found to be a negative prognostic sign in all treatments.

Blatt and Gohar (2004) recently returned to the original Menninger data and investigated whether by subdividing the patients along a new set of diagnostic lines, it would be possible to parse out the differential benefit of psychoanalysis and psychodynamic psychotherapy. They found that with more ruminative, self-reflective patients (termed by Blatt “introjective”), psychoanalysis appeared to have a greater benefit by fostering adaptive interpersonal capacities and reduce maladaptive interpersonal tendencies. In more dependent, unreflective, affectively labile individuals (termed by Blatt “anaclitic”), less intensive supportive-expressive therapy was more effective, possibly by containing or limiting their associative capacities (Blatt & Shahar, 2004).

With major outcome studies begun in the 1950’s already occupying many of the centres of psychoanalysis, particularly in the United States, the 1960’s and 1970’s saw few new large-scale studies, but a diversification of the sites where research was being conducted and involvement of new patient populations. In addition, having witnessed the obstacles faced by large studies, researchers began designing studies with narrower and more specific goals. Studies at the Berlin Central Institute for Psychogenic Illnesses, by Barendregt and colleagues in Amsterdam, Cremerius in Germany, and by Bieber (1962), O’Connor (1964), and Zetzel (1968) in the United States, looked at more specific populations and asked how psychoanalysis affected a range of variables, ranging from the utilization of medical services to projective tests (Fisher & Greenberg, 1977; Fonagy et al., 2001). Also during this period, two large scale studies got underway at the New York Psychoanalytic Association (Erle, 1979; 1984) and the Boston Psychoanalytic Institute (Kantrowitz, 1993).

Since 1980 the empirical study of psychoanalytic outcome has shifted out of the United States, and into Europe and Latin America. Of the 18 research projects that began during this time period, seven were based entirely in Germany, four in England, three in Scandinavia, one in Holland, one in Latin America, one as a joint European effort, and one in the United States (although it ended before producing any results). Most notable of these are projects at the University of Heidelberg Psychosomatic Clinic (Kordy, von Rad, & Senf, 1983, 1988, 1989; Senf, Kordy, von Rad, & Bräutigam, 1984), by Rudolf and colleagues in Berlin (Rudolf, 1991), and by Sandell and colleagues at the Stockholm Karolinska Psychoanalytic Institute (Sandell, 1987a; 1987b; 1987c; 1988; 1991; 2000).

Because of its innovative design and interesting findings, the Stockholm study warrants detailed description. Sandell and colleagues at the Stockholm Karolinska Psychoanalytic Institute carried out a quasi-experimental study of government subsidized psychoanalysis in which they collected data on three occasions (1994, 1995, and 1996, referred to as a “three-wave panel design”) from a large sample at various stages of therapy (Fonagy et al., 2001; Grant & Sandell, 2004; Sandell, 1987a; 1987b; 1987c; 1988; 1991; Sandell et al., 2000). In total, data were collected from 60 patients in psychoanalysis, 140 patients in long-term psychodynamic psychotherapy, 500 patients on the waiting list for subsidized therapy, 400 subjects from a random community sample, and 250 university students. Patients were assigned to psychoanalysis or psychotherapy on the basis of a clinical assessment of the patient’s need of and suitability for each type of treatment. At each time point, self-report data regarding psychiatric symptomatology, social adjustment, and “sense of coherence” were collected. At two time points (1994 and 1995), qualitative interviews were performed with a subsample of 20 patients who had completed psychoanalysis, 20 patients who had completed psychotherapy, and 20 patients who were still on the waiting list. These interviews were rated according to

Sandell's Change after Psychotherapy Scales (CHAP) which quantify change along four dimensions: symptoms, adaptive capacity, self-insight, and basic conflicts (the last of which is offered as a rating of "structural change") (1987b). Data on work absenteeism and health care utilization were collected for all patients in the study from the national health insurance and health care authorities (Sandell et al., 2000).

In 1996, all 316 analysts and therapists in the study, as well as 350 licensed Swedish therapists not in the study, were asked to fill out a "therapeutic identity" questionnaire which asked about (1) basic education and professional training, (2) professional experience, (3) personal/training therapy/ies, (4) theoretical orientation, (5) therapeutic ideals and technical approach, and (6) ideas about the nature of psychotherapy and the human mind. Completed surveys were returned by 209 analysts and therapists in the study (Sandell et al., 2000).

A variety of factors were found to differentiate patients in psychoanalysis and psychotherapy. Analytic patients were older, more frequently male, and more likely to have been married, to have children, and to hold university degrees. They were also more likely to have previously utilized psychotherapy services, whereas psychotherapy patients tended to have utilized inpatient or outpatient clinics and emergency rooms. Following the treatment being studied, psychotherapy patients were twice as likely as analytic patients to seek further psychotherapeutic treatment. The two samples did not differ significantly in any diagnostic categories or severity of disturbance (Sandell et al., 2000).

Findings revealed that both psychotherapy and analytic patients improved steadily during treatment on the symptom and sense of coherence measures, but that after treatment was concluded analytic cases continued to improve over the next three years, while the change in psychotherapy cases leveled off. The overall social adjustment scale showed similar improvement in analytic and psychotherapy groups. Most of the subscales of the social adjustment measure (all except for work and friends scales)

showed an initial deterioration during treatment, but then rebounded. The greatest improvement was in the work subscale. When lying in the non-clinical range on all three outcome measures was taken as the criteria for psychological health, the analytic group improved from 12% healthy to over 70% healthy between pre-treatment and three year post-termination, whereas the psychotherapy group went from 30% to 55%. In comparison, 84% of subjects in the combined norm group met this criterion. The CHAP scales confirmed the finding of general improvement, with the greatest changes observed in symptoms and adaptive capacity. Analytic patients were found to improve significantly more than psychotherapy patients on ratings of basic conflict resolution, supporting the hope that this would be a measure of “structural change” specific to the analytic situation (Sandell et al., 2000).

Several interesting findings resulted from the study of therapist factors and their association with outcome in the various treatment groups. Older therapists achieved the best outcomes, irrespective of therapist or patient gender and type of therapy. The second youngest group of therapists, not the youngest, had the worst outcomes overall. Time practicing therapy was positively related to outcome, but only the time after licensing, not the time before. Since patient assignment was non-random, it is not clear whether these effects resulted from differences in therapeutic skills or of healthier patients selecting more experienced therapists. Psychoanalysts did tend to be older and have greater years of experience than psychotherapists, but this difference did not entirely account for the increased efficacy of psychoanalysis. Psychoanalytic training was not associated with better outcomes in patients treated with psychotherapy (Sandell et al., 2000).

Path analyses performed on 156 patients who had already terminated their treatments by the first assessment revealed further findings on the effects of session frequency, duration, and subsidization on outcome. Long duration and high frequency

were positively associated with outcome only when they occurred in conjunction; long psychotherapies and short psychoanalyses on average were less effective than long psychoanalyses and short psychotherapies. This effect was significant only on the basis of the third follow-up, three years after termination. Subsidization was observed to have no effect on outcome (Sandell et al., 2000).

Changes in health care utilization and work absenteeism were analyzed separately for psychotherapy and psychoanalytic samples. In the psychotherapy group, utilization of social welfare and somatic outpatient services and work absenteeism dropped both during and after treatment. In the psychoanalytic group, despite an improvement in self-rated general health and capacity for work, there was a significantly increased dependence on social welfare, use of somatic outpatient services, consumption of psychoactive medication, and work absenteeism (Sandell et al., 2000).

A factor analysis of therapist-rated items about therapeutic style and beliefs followed by a cluster analysis, yielded three groups of therapists or analysts: “classical psychoanalytic,” and two eclectic groups, “supportive-interpersonal,” and “supportive-intuitive.” The “classical psychoanalytic” group was associated, but did not completely overlap, with those who had undergone psychoanalytic training. Among psychoanalytic cases, patients did equally well with analysts in the “classical” or “eclectic” groups. However, among psychotherapy cases, patients did significantly better with “eclectic” than with “classical psychoanalytic” therapists. In particular, therapists who were high on the factors “kindness,” “self-disclosure,” “supportiveness”, “insight,” “neutrality”, and “art” (he or she considered psychotherapy more a work of art than a craft or science) were more successful in psychotherapy (Sandell et al., 2000).

Additional psychoanalytic research studies are being conducted at the German Psychoanalytic Association (Leuzinger-Bohleber, 1999), by Huber and Klug in Munich (2004), by members of the European Multi-site Collaborative Study (also called AHMOS

for its sites: Amsterdam, Helsinki, Milan, Oslo, and Stockholm, and by Fonagy and colleagues at the Anna Freud Centre in London (Fonagy et al., 2001). Studies of psychodynamic psychotherapy and psychoanalytic process only are too numerous to review here and will be summarized in later chapters where relevant.

1.4 Current answers to major questions about psychoanalysis from outcome research

Given the multitude of studies listed above, it would be useful to review briefly how, as a body of research, they answer some of the basic questions addressed. Due to the complexity of the questions asked and the problems in research methodology, it would be hasty to presume that these studies reach a consensus on all or even most questions asked. However, there do appear to be certain directions in which the findings point, that will be helpful to identify before designing future projects.

1.4.1 Is psychoanalysis effective?

Virtually all the psychoanalytic outcome studies performed over the past 87 years suggest that a majority of patients in psychoanalysis do appear to improve in association with their treatment. Rates of improvement, depending on how the data were collected and how the criteria for “improvement” were set, range between 60 and 90 percent. In the studies that collected outcome data via analyst report, the reported improvements were on the higher end of this range: Coriat (1917), the American Psychoanalytic survey (Hamburg et al., 1967), and the Columbia Record Project (Bachrach, 1993) cite rates of 73%, 97%, and 88-100%, respectively. In Europe, the Berlin (Rudolf, 1991) and Stuttgart (Fonagy et al., 2001) studies queried analysts about structural developments and achievement of treatment goals and report significant improvement in 83% and 67% of their respective samples. When improvement was gauged by psychological testing or an independent rating of patient information, rates were slightly lower. The Menninger PRP

(Wallerstein, 1986) reported improvement in 59% of patients and an average increase on their 100-point global functioning scale of 14 points. The component of the Columbia Records Project (Bachrach, 1993) using independent raters reported improvement rates of 56-91%, the early Boston Psychoanalytic project (Sashin et al., 1975) 67%, the New York project (Erle, 1979) 60%, the Heidelberg study (Kordy et al., 1983) 87%, and the Anna Freud Centre chart-review of children 72% (Fonagy & Target, 1996). Newer studies, such as Sandell (2000) and the Berlin Jungian (Fonagy et al., 2001) project report improvement on a wide range of measures, including self-report, social adjustment, and “adaptive capacity.” Research underway by the European Multi-site Collaborative and the Latin American groups (Fonagy et al., 2001) is likely to shed more light on the subject of overall rates of improvement.

A helpful validity check by many of these studies has been to show that a “dose-response” relationship exists between psychoanalysis and improvement. Howard and colleagues (Howard, Kopta, Krause, & Orlinsky, 1986) argued that being able to show that longer treatments are associated with better outcome makes it more likely that treatment and outcome are causally linked. Both analyses of the Columbia records project (Bachrach, 1993), the New York Psychoanalytic study (Erle & Goldberg, 1984), and the Berlin Jungian project (Fonagy et al., 2001) reported these relationships. Kordy and colleagues in Heidelberg (2001) found a dose-response relationship with an optimal effect at 2.5 years or 160 sessions. In the Anna Freud Centre chart-review of child psychoanalysis and psychotherapy, independent associations of duration and frequency of treatment to outcome were found, whereas Sandell (2000) found an association of duration with outcome only *in conjunction* with high frequency treatments. Kantrowitz (1993) is one of the few investigators to report the conspicuous absence of a relationship between treatment length and outcome.

In a related vein, a few studies have shown that treatments identified as “completed” by analysts, and terminated by mutual agreement, are associated with greater benefit for the patient. The original Boston Psychoanalytic study (Sashin et al., 1975), both Columbia projects (Bachrach, 1993) and the Berlin study (Fonagy et al., 2001; Rudolf, 1991) demonstrated this effect.

Throughout the history of psychoanalytic research, investigators have been thoughtful in presenting multiple ways of viewing improvement. The Stockholm (Sandell et al., 2000), Stuttgart (Fonagy et al., 2001), German Psychoanalytic (Leuzinger-Bohleber, 1999), and Berlin Jungian studies (Fonagy et al., 2001), using a variety of outcome measures, have found that at several years follow-up, the functioning and well-being of a group treated with psychoanalysis is comparable to that of the non-clinical population. The latter three of these groups have also demonstrated that patients have an improved capacity to work and reduced absenteeism after analysis. Preliminary results of the Anna Freud Centre follow-up study have revealed that subjects who are rated from their charts as having “good treatment outcome” demonstrate work functioning as adults that is as good as untreated siblings or controls (Fonagy et al., 2001).

Berlin Jungian and German Psychoanalytic groups have also demonstrated that patients treated in psychoanalysis make fewer health insurance claims, both for psychotropic medications and for non-mental health medical services (Fonagy et al., 2001; Leuzinger-Bohleber, 1999). The only contrary finding emerges from Sandell’s study (2000) in which he finds that analytic patients showed *increased* health care utilization and work absenteeism after completing treatment, despite the fact that they rate their own general health and capacity for work as *higher* than when they started treatment. Sandell suggests that these patients were underutilizing health care and taking fewer days off than a normal sample when they started therapy, and thus learned through analysis to take

proper advantage of these services. Comparison of the overall rates for health care utilization and work absenteeism in the various groups supports this assertion.

Another reassuring finding from many psychoanalytic outcome studies has been that the improvements measured immediately after treatment persist for years into the future. Cremerius (Fisher & Greenberg, 1977) noted maintained improvement 8 to 10 years after termination, the Heidelberg study (Kordy et al., 1983) at 3.5 years, the Stuttgart study (Fonagy et al., 2001) at 4 years, and the Berlin Jungian study (Fonagy et al., 2001) at 6 years. The Boston follow-up study (Kantrowitz, 1993), the Stockholm study (Sandell et al., 2000), and Heinicke's study of children (1986) all note continued improvement between termination and follow-up, though, as Kantrowitz points out, they are unable to predict who will continue to change as opposed to who will remain the same. Investigators in the AFC follow-up study make the interesting observation that subjects who received successful analytic treatment as children have better love relationships, are more likely to be secure on Main's Adult Attachment Interview (AAI), have higher mentalizing capacity, and experience fewer adverse life events than do their untreated siblings *20 years* after the completion of treatment (Fonagy et al., 2001).

The optimistic results of psychoanalytic outcome studies on the question of treatment efficacy must be accompanied by two important caveats. First, as discussed in the previous section, the method for quantifying outcome is highly questionable in most studies, and often involves the potential for significant bias from impossible-to-blind patient or clinician reporters. Second, raw statistics about the percentage of patients who improve from the beginning to end of therapy tell us little about what caused the improvement. In a classic critique of psychoanalysis, Eysenck (1952) asserted that all the change reported by early analytic outcome researchers could be accounted for by the natural course of mental illness. Although, many investigators have refuted Eysenck's statistics, claiming that he over-estimated the incidence of spontaneous regression, and

the methodology for measuring outcome has improved considerably since the papers he critiqued, the fact remains that without a proper control or process measures there is little to causally link treatment to outcome. One concerning result in this regard has been that in at least four studies, although patients improved symptomatically, investigators were surprised to record a very low rate of “analytic success” or “analytic process” as measured by analysts or independent raters. The American Psychoanalytic survey (Hamburg et al., 1967) reported that only 27% of patients achieved a “symptom cure,” the Boston follow-up study (Kantrowitz, 1993) showed a “successful analytic result” in 40% of patients, the New York Psychoanalytic outcome study (Erle, 1979) measured a 17% rate of “analytic process” in one sample and 25% in another, and the Columbia Records study (Bachrach, 1993) noted “analytic process” in only 43 to 50% of patients. Clearly, to understand more about the rates of efficacy of psychoanalysis, these numbers need to be reconciled with those regarding symptomatic improvement.

1.4.2. How do psychoanalysis and psychotherapy differ?

The majority of psychoanalytic outcome studies have suggested that subjects undergoing psychoanalysis improve more than those undergoing less frequent psychodynamic psychotherapy. In none of the following studies were subjects randomly assigned to psychoanalytic and psychotherapy treatment groups. The American Psychoanalytic survey (Hamburg et al., 1967) reported psychoanalysis to be superior across all diagnostic groups, but acknowledged that psychoanalytic patients were on average higher functioning than their counterparts in psychotherapy. Both analyses of the Columbia Records project (Bachrach, 1993) found that analytic patients did better than general psychotherapy and psychodynamic psychotherapy patients, though in the first study clinic analytic patients (as opposed to private analytic patients) did worst of all. The Berlin study (Fonagy et al., 2001; Rudolf, 1991) found analysis to be superior to both outpatient and inpatient dynamic psychotherapy. In Heidelberg (Kordy et al., 1983),

analysis and psychotherapy did equally well in improving symptomatology, but significantly more patients in psychoanalysis met individual treatment goals (72% vs. 50%), as defined by independent raters at the outset of treatment. In Stockholm (Grant & Sandell, 2004; Sandell et al., 2000) analytic and psychotherapy patients improved similarly during therapy, but while improvement levelled off immediately following psychotherapy, patients who had been in psychoanalysis continued to improve three years beyond termination. Analytic patients were also less likely to seek further psychotherapy. The AFC chart-review project (Fonagy & Target, 1996) found frequency of therapy to be an independent predictor of improvement in children, and analysis was particularly beneficial in children with severe disorders and those 12 years old and younger (older children did equally well with psychotherapy).

To date, the only completed randomized comparison of psychoanalysis and psychotherapy is Heinicke's (1965) study of the treatment of reading disorders in children, in which he found that the benefits of analysis were greater than those in psychotherapy and these were continued and extended in follow-up. All other studies have been faced with the possibility that subjects appear to do better in psychoanalysis because they are self- or analyst-selected for reasons that predict improvement. There is evidence that patients who self-select for psychoanalysis are older, better educated, more likely male, and have had more previous psychotherapy than those selecting psychotherapy (Sandell et al., 2000). It has been reassuring, though, that most studies have been unable to predict improvement on a wide range of initial variables (as described below), making it surprising that patients would be able to consistently differentiate themselves, or analysts would be able to consistently select their patients, along these lines.

A number of ongoing studies promise to shed light on the differences in efficacy between psychoanalysis and psychotherapy. European Multi-site Collaborative, Latin

American, and AFC follow-up studies will continue to look at this question without randomization. Meanwhile, Huber and Klug in Munich (Fonagy et al., 2001; Huber & Klug, 2004) are carrying out the first near-randomized (due to small sample size some stratification and group matching is being done) comparison of psychoanalysis and psychotherapy in adults, in the treatment of depression. The AFC prospective study will be the first large scale randomized comparison of psychoanalysis and psychotherapy in children. Several ongoing studies, including the Helsinki site of the European Multi-site Collaborative and the Cornell-Westchester prospective trial are applying randomization to comparison of psychodynamic psychotherapy and non-psychodynamic psychotherapy, but have not extended this to include psychoanalysis proper (Fonagy et al., 2001).

Outcome studies have presented other interesting ways to distinguish the outcome of psychoanalysis and psychotherapy. The American Psychoanalytic survey (Hamburg et al., 1967), the first analysis of the Columbia Records Project (Bachrach, 1993), and the Berlin study (Rudolf, 1991) showed that patients in psychoanalysis are two to three times more likely to complete their therapy and terminate under mutual agreement than patients in psychotherapy. Heidelberg (Kordy et al., 1983) and Berlin (Rudolf, 1991) studies found that psychoanalysis is better for psychosomatic disorders and symptoms than psychotherapy. The AFC chart-review (Fonagy & Target, 1996) found psychoanalysis more beneficial than psychotherapy particularly for children with emotional (as opposed to conduct) disorders. Studies are currently underway in Munich and Helsinki to test the differential effects of these therapies on patients with depression and anxiety disorders, with the hypothesis that psychoanalysis will have greater extent and durability of improvement.

A number of contemporary studies are investigating the question of whether the nature of change is different in psychoanalysis and psychotherapy. Sandell (2000) has found that analytic patients change more on ratings of basic conflict resolution than do

psychotherapy patients. Leuzinger-Bohleber (1999) is investigating this question using Wallerstein's Scales of Psychological Capacities, the Heidelberg-Berlin study using Stiles's APES measure, and the European Multi-Site Collaborative, Latin American, and Munich projects (Fonagy et al., 2001) with broad batteries of measures. A few of the European Collaborative sites, including Helsinki, are particularly interested in the question whether analytic patients will show greater improvement on Fonagy's Reflective Function measure, indicating an improved capacity to mentalize, than patients in psychotherapy (Fonagy et al., 2001).

A smaller, but significant, number of studies have suggested that the differences between psychoanalytic and psychotherapy outcome are not as great as originally suspected. Most prominently, the Menninger PRP found near identical improvement rates of 59% and 60% in analytic and psychotherapy patients, respectively. Wallerstein (1986) suggested that this was related to the severity of illness in the sample, the high rate of supportive interventions in both groups, and the lack of an exclusive link between an interpretive style and structural change. However, this conclusion is called into question by Kernberg's, and much more recently, Blatt's analysis of the Menninger data suggesting that real distinctions are possible, once the patients are more carefully characterized (Blatt & Shahar, 2004; Kernberg, 1973).

Surveys by Heilbrunn (1966) and Cartwright (1966) also found little difference between psychotherapy and psychoanalysis. In Heidelberg, Kordy (1983) noted that although analytic patients were still moderately better off than psychotherapy patients 3.5 years after follow-up, this difference had diminished since termination and might continue to do so. He also observed that patients reported being more satisfied with therapy than analysis (52% vs. 16%, respectively, indicated they were "very satisfied"). Finally, Sandell (2000) was surprised to find that analytic patients increased their utilization of social welfare and outpatient medical services whereas psychotherapy

patients decreased utilization of these services. As described above, he believes this is due to correction of underutilization on the part of analytic patients.

1.4.3. What patients are more likely to benefit from psychoanalysis?

One of the most frustrating areas of psychoanalytic outcome research has been the struggle over finding ways to predict who would benefit from psychoanalysis. The motivation for most of the clinicians and many of the researchers who pursue outcome research has been to determine criteria that would allow them to recommend psychoanalysis specifically for those people who would benefit from it. Particularly in the modern era of limited resources and focus on the cost-effectiveness of psychoanalysis, the holy grail of outcome research has become to outline a group of people for whom psychoanalysis is the most effective and most efficient form of treatment. Sadly, the vast majority of these attempts has met with failure.

Very early studies of psychoanalysis, performed before the advent of neuroleptics and contemporary neurobiological theories of psychosis, were instructive in demonstrating that psychoanalysis is far more effective in neurotic than in psychotic patients. Fenichel (1930), Jones (1936), Alexander (1937), Knight (1941), and the American Psychoanalytic survey (Hamburg et al., 1967) showed that neurotic patients improved at rates as much as seven times higher than psychotic patients. However, among the various diagnostic categories subsumed under “non-psychotic” patients, it was impossible to forecast outcome. Coriat (1917) found equivalent outcome across 14 diagnostic categories. Using ever more sophisticated and comprehensive techniques of collecting diagnostic information, both Boston (Kantrowitz, 1993; Sashin et al., 1975), both Columbia (Bachrach, 1993) and New York (Erle, 1979) projects were unable to predict outcome from initial diagnostic information. Most studies report a similar lack of findings using demographic information as predictors. These results led Waelder to lament, “This state of affairs is not just due to the youth of our science or to the alleged failure to adhere to

rigid standards of investigation and verification, but to [quantitative] factors which appear to be inherent in the subject and which therefore are, on the whole, unalterable” (Waelder, 1963).

Although they are relatively rare and unreproduced, the few findings linking initial evaluation to outcome are worth reviewing. In his quantitative analysis of the Menninger PRP, Kernberg (1972) found that patients with high ego strength and high quality of interpersonal relationships do better in psychoanalysis and psychotherapy. Blatt and Shohar (2004) retrospectively identified a population of “introjective” patients (ruminative, self-reflective, and suffering from poor interpersonal relatedness) who do better in psychoanalysis. Fonagy and Target (1996) found that diagnoses in children were significantly associated with outcome from treatment. Phobias and anxiety disorders were most likely to change, and obsessive compulsive disorder and depressive disorders least likely. In general, children with disruptive disorders did not benefit as much as children with emotional disorders, often terminating prematurely, and the presence of a comorbid anxiety disorder was a positive prognostic sign in disruptive children. On demographics, the American Psychoanalytic survey (Hamburg et al., 1967) found that depressed men benefited more from analysis than depressed women, but schizophrenic women did better than men. Wallerstein (1986) observed that women improved more than men, but postulated that this was due to a lower rate of substance abuse and less frequent referral for “heroic indications.” The only study to link socioeconomic status to improvement in psychoanalysis, the Columbia Records project (Bachrach, 1993), noted that patients seen privately improved more often and were more likely to terminate mutually than patients seen through a low fee clinic.

In response to the difficulty using standard diagnostic and demographic information to predict outcome, a number of theoretical and empirical analytic papers began to focus on an operationalisation of “analyzability.” When analysts or independent

raters were asked to assess analyzability retrospectively based on information from an entire treatment, not surprisingly, these judgments were associated with outcome (Bachrach, 1993). However, when they were made prospectively, no relation of analyzability and outcome was observed (Erle, 1979; 1984).

The study of links between initial diagnosis and outcome have been somewhat more successful in studies of psychodynamic psychotherapy, possibly due to the larger sample sizes and greater methodological flexibility in these studies. In Norway, Monsen and colleagues (Monsen, Odland, Faugli, Daae, & Eilertsen, 1995a; 1995b) have shown that personality disordered patients are less likely to improve than non-personality disordered subjects matched for severity of illness. Najavits and Gunderson (1995) have found that borderline patients with comorbid anxiety disorders do worse than those without anxiety disorders. Several process studies have shown that patients can be differentiated on the basis of object relation patterns (Blatt, 1992; Davies-Osterkamp, Strauss, & Schmitz, 1996; Horowitz, Marmar, Weiss, Kaltreider, & Wilner, 1986) as to who will respond best to psychotherapy.

1.4.4. What makes psychoanalysis effective?

The study of how analytic treatment parameters are related to outcome has been moderately more successful than that of initial conditions. Most significant are the findings, described in parts 1 and 2 of this section, that longer and more frequent (i.e., psychoanalysis vs. psychotherapy) therapy is more likely to be completed and more likely to lead to patient improvement. It is less clear, though, whether “analytic process” is significantly related to outcome. New York (Erle, 1979) and Columbia (Bachrach, 1993) projects failed to show a consistent relationship. At the very least they show that development of “transference neurosis” as defined by their raters is neither necessary nor sufficient for producing therapeutic change. Kernberg (1972) and Sandell (2000) describe more complicated connections. From his analysis of the Menninger PRP, Kernberg

concluded that treatment of severe personality disorders requires a combination of supportive and expressive (i.e., interpretive) techniques, a focus on the transference, and a “high-skill” therapist. Sandell studied the belief systems of therapists and found that psychoanalytic patients do equally well with “classical” and “eclectic” treaters, while psychotherapy patients do better with “eclectic” treaters (i.e., rate themselves higher on kindness, self-disclosure, supportiveness, insight, neutrality, and thinking of treatment as more of an art than a science).

Several researchers have emphasized the alliance or fit between patient and therapist as the most significant predictor of success. In Horwitz’s analysis of the Menninger PRP (1974) alliance was the best predictor of outcome. In Berlin, Rudolf (1991) found that therapist assessment of alliance was a strong predictor of outcome, and patient assessment of alliance somewhat weaker but still significant. Kantrowitz (1993) has written the most extensively and elegantly about the way in which interdigitation of patient and analyst styles predicts improvement.

A few aspects of the analytic process have not been adequately studied and may have an important effect on outcome that accounts for variation among studies. Sandell (2000) reports that therapist experience has a “J” shaped relationship with analytic outcome. In the few years after licensing the therapist’s success rate decreased with age (possibly because of increasing commitments or taking more difficult patients) but then begins to rise again, and is best for the most senior analysts. Bachrach (1993) observed that less experienced analysts were more likely to have patients with premature termination and less improvement. Despite these findings, many studies continue to use a heavy proportion of inexperienced analysts in their studies without a way to account for the effect of this selection on their findings. Another inadequately studied effect is the relationship between the amount that patients pay to their outcome. Sandell (2000) found no such effect, but further research is required.

One of the most exciting areas of current research in psychoanalysis is the measurement of how an improvement of the patient's capacity for mentalization is related to therapeutic improvement. The European Multi-Site Collaborative study (2001) is using Fonagy and Target's reflective function (RF) scale, Main's AAI classification system, and the monthly Periodical Rating Scale process measure to attempt to link relationship and cognitive styles with symptoms and improvement. The AFC follow-up study (Fonagy et al., 2001) has suggested that traumatized children whose treatment involved elaboration of mental states are doing better at 20 year follow-up than those whose treatment focused on interpretations of conflict.

1.4.5. What is the process of psychoanalysis?

A satisfying understanding of psychoanalytic outcome would involve not only the ability to predict who would improve in what types of therapy, but would also reveal the causal mechanisms that connect these variables through the analytic process to change. Historically, most outcome studies were able to shed light on this question only peripherally. The greatest promise today lies in specific studies of analytic process, reviewed in Chapter 2, many of which do not attempt to tackle the problems of large scale outcome studies. A few interesting findings have appeared, though, in outcome studies.

Early on, in the American Psychoanalytic survey (Hamburg et al., 1967), the Columbia Records Project (Bachrach, 1993), and the New York Psychoanalytic study (Erle, 1979) it was noted with disappointment that fewer than expected analyses terminated by mutual agreement or were considered "complete" by analysts or independent judges. Rates for "successful resolution of the analytic process" ranged between 57% and 66% in these studies, and were considerably lower than the rates of symptomatic improvement. A related finding in the New York study and several studies using the Pfeffer outcome method (Norman, Blacker, Oremland, & Barrett, 1976;

Oremland et al., 1975; Schlessinger & Robbins, 1974), was that transference neurosis was not obliterated in even successful analytic cases, refuting the theoretical assumption that successful resolution of the transference is a necessary part of a successful treatment.

Another surprising finding from several studies was that no matter how thorough the initial attempt at diagnosis, many patients revealed greater pathology during the treatment than was initially suspected. This may be postulated to result either from the failure of initial diagnostic techniques or the regression that a successful analytic treatment brings about in many patients. Wallerstein (1986), Erle (1979), and Kantrowitz (1993) suggest that such a pattern in their studies contributed to the difficulty in predicting outcome from initial diagnosis. A potentially related finding from Sandell (2000) is that several of the subscales of the social adjustment measure initially deteriorated after the onset of therapy before rebounding and showing improvement by termination. This observation has probably not been reported by other studies because of the scarcity of frequent repeated measures of functioning during treatment. As process and process-outcome studies collect further data, many details of the day to day course of analysis will be illuminated. For example, the Latin American study plans to focus on the way in which the very act of conducting research affects patients and analysts within their study (Fonagy et al., 2001).

1.4.6. What do we know about psychoanalytic treatment of young adults?

For the purposes of the research project described in this thesis which studies the treatment of young adults, it is relevant to consider the available outcome data that is specific to this population. Although the consideration of young adults has not been an explicit question raised by any of the outcome studies reviewed above, it emerges that it is an important component to understanding these studies, as most had samples in which young adults were heavily represented. Given this, it is surprising that no investigators

linked the age of their sample and the known theoretical issues in dealing with such a sample to their results.

The period of young adulthood is inherently difficult to define, but is usually considered to start at the end of adolescence, anywhere from 18 years old to the early twenties, and proceeding to the onset of mature adulthood in the mid to late twenties. Many analysts, starting with Freud, have discussed treatment of late adolescents and young adults without describing their problems as distinct from adolescents or mature adults. In more recent times, analysts have begun to ask questions specific to this time period. Blos (1977) described four developmental tasks of adolescence that define the early challenges of young adulthood: (1) negotiation of the second individuation process, (2) ego continuity, (3) relative mastery of the inevitable accumulation of traumas during infancy, childhood, and adolescence, (4) establishing sexual identity. Jacobs (1987) added to this list (5) movement from narcissistic self-involvement of adolescence to more intimate and enduring object relations, (6) development of greater breadth of ego functions less governed by conflict and more energy for new interests and activities, and (7) reorganization of superego and ego ideal.

The analyst of a young adult must face several practical consequences of the developmental stage and life circumstances. First, because of the important changes taking place, it is difficult for the analyst, and the outcome researcher, to tell the effect of analysis from natural maturation. Similarly it may be hard to distinguish developmental conflict from psychopathology (Jacobs & Chuscd, 1987). Next, because of the interaction of the analysis and developmental issues, the young adult often experiences regression in the analysis as an extension of their own psychic reorganization and may become resistant to this, declaring "I need to get on with my life." As a result it may be particularly difficult and not necessarily advantageous for the young adult to develop a full transference neurosis (Adatto, 1980). Classical writers still regard the failure to

develop a transference neurosis in young adults as defensive (Blos, 1967) or as a failure of oedipal resolution (Calef in Escoll, 1987). Scharfman (Jacobs & Chused, 1987) suggests that the young adult's transient idealization and then deidealization of the analyst is a normal part of the treatment and should not be interpreted solely as a negative transference. In particular he feels that the homosexual component of the idealized relationship cannot be fully analysed. He states that the "transference neurosis is not organized with the same clarity and intensity in the young adult" (Jacobs & Chused, 1987, p. 178). Furthermore, the reason that many individuals analysed in late adolescence and young adulthood pursue careers related to psychoanalysis, may be because of an ongoing unresolved identification with the analyst that goes beyond the analyzing function (though many also began analysis because of a conscious career choice).

In addition to facing the maturational challenges of contemporary society, young adults are in a very real situation of transition. Young adult patients often move frequently, sometimes on school holidays or because of job changes. They either face the challenge of negotiating with their parents to pay for sessions, or may request a period of less frequent sessions because of their own financial difficulties. As a result of these changes, it might seem that young adults are more suitable for the flexible schedule of psychotherapy, but several have argued that putting off potentially life-altering analysis at this highly important stage is often not recommended (Escoll, 1987).

Given the serious theoretical and practical issues that make analysis of young adults different from the prototype conceived for mature adults, it is surprising how many outcome studies use samples in which young adults are heavily represented. This is no doubt related to the reasons for using inexperienced analysts; young adults are more likely to be in need of reduced fee therapy supplied by research studies and trainees. Of Freud's 12 "major" cases, as many as nine may have been between the ages of 18 and 30

during treatment, including the well known Wolf Man and Rat Man. Of 107 “minor” cases, 22 were between the ages of 21 and 30 (Brody, 1970). Many early analytic surveys did not report frequencies within age ranges but undoubtedly contained a large number of young adults. In the early Boston study (Knapp et al., 1960), the age range of 100 subjects was 20 to 41 and the mean 27. Siegel (1962), Weintraub and Aronson (1968), and Karush (1956) all report a high percentage of their patients in the 21 to 30 range.

Major studies of psychoanalytic outcome consistently report high representation of young adults. One-quarter to one-half of the 42 Menninger PRP patients were between 18 and 30 (mean age of the sample was 31), and several were identified with “unresolved adolescent problems as central psychopathological issues” (Wallerstein, 1986, p. 522). Between one-quarter and one-third of patients in the American Psychoanalytic survey (Hamburg et al., 1967) and the Columbia Records Project (Weber, Solomon et al., 1985) were young adults, and many of the case studies studied in follow-up by Pfeffer, Oremland, and colleagues (Oremland et al., 1975). The Heidelberg and Berlin Jungian studies both report mean ages in the early 30’s.

Most strikingly, 35 of the 40 patients in the initial New York Psychoanalytic sample (Erle, 1979) were between 21 and 28. Interestingly, this study was notable, relative to other similar studies, for a low improvement rate (60%), low rate of “analytic process” (17%), and low rate of patients deemed “analyzable with resolution of the transference” (9%). In almost half, the diagnosis at termination was worse than at intake. The New York comparison group, drawn from somewhat older patients seen by graduate analysts (as opposed to candidate analysts) had a higher rate of improvement and analytic process. In the Boston follow-up study (Kantrowitz, Katz, Paolitto, Sashin, & Solomon, 1987), all 22 patients were between 21 and 32 at initial evaluation, yet this is not discussed relative to absence of predictors for outcome.

The situation regarding consideration of age and psychoanalytic outcome does not appear to be improving. Several otherwise good contemporary studies (Leuzinger-Bohleber, 1999; Rudolf, 1991; Sandell et al., 2000) have so far not reported the ages of their subjects in early publications. Ongoing studies such as Heidelberg-Berlin, Munich Depression, Latin American, and the European Multi-Site Collaborative studies (Fonagy et al., 2001) do not address issues of age in their design. Clearly, this is an area that will require greater attention in the future.

1.5 Conclusion

In the above chapter we have presented a brief review of 87 years of psychoanalytic outcome research, a set of questions raised by those studies regarding methods of evaluating patient diagnosis, treatment, and outcome, and a summary of the findings of those studies through the lens of the most relevant questions posed by outcome research. It is hardly necessary to point out that an enormous amount of work has been done in this field and a great number of interesting findings uncovered. However, it is difficult to leave a review of psychoanalytic outcome research with the feeling that any of the questions posed has received anything near to a firm or final answer. Sadly, it also seems that the path toward better studies has been even more indirect than is expected in scientific research. Many of the problems with outcome research seem to have been clear to investigators from the first such projects, yet succeeding generations of researchers have had to rediscover them as opposed to using the experience of their forebears to move forward. A fitting, yet ironic, summary of this process is provided by Merton Gill (1979) and quoted by Wallerstein in his introduction to the Menninger PRP report (1986). Gill's comments come at the end of an article defending the centrality of interpreting transference in the here and now to psychoanalytic technique.

I close with a statement of a conviction designed to set this paper into a broader perspective of psychoanalytic theory and research. The points I have made are not new. They are present in varying degrees of clarity and emphasis throughout our literature. But like so many other aspects of psychoanalytic theory and practice, they fade in and out of prominence and are rediscovered again and again, possibly occasionally with some modest conceptual advance, but often with a newness attributable only to ignorance of past contributions. There are doubtless many reasons for this phenomenon. But not the least, in my opinion, is the almost total absence of systematic and controlled research in the psychoanalytic situation. I mean such research in contrast to the customary clinical research. I believe that only with such systematic and controlled research will analytic findings become solid and secure knowledge instead of being subject to erosion again and again by waves of fashion and to what Ernst Lewy (1941) long ago called the "return of the repression" to designate the retreat by psychoanalysts from insights they had once reached (Gill, 1979).

The very attempts at systematic and controlled research in which Gill concentrates his hope for avoiding the cycle of forgetting and rediscovery of psychoanalytic principles, have been subject to these same frustrating trends. We hope that future studies find a way to break this cycle.

CHAPTER 2. PSYCHOTHERAPY RESEARCH METHODOLOGY

2.1 Introduction

In order to study psychoanalytic outcome research in the context of empirical science and to think systematically about advancing its methodology, we must review the fundamentals of psychotherapy research in general. Starting in the 1940's, non-psychoanalytic forms of psychotherapy were introduced and investigated by researchers seeking to demonstrate their effectiveness. It was quickly apparent that some of the new forms of therapy were more amenable to research than psychoanalysis as their goals were more concretely defined and the length of therapy shorter. Due to a combination of practical and cultural factors, non-psychodynamic psychotherapy research accelerated at a rapid rate over the following decades. The pace of research was fueled in part by the growth of the number of different forms and orientations of psychotherapy being practised. In the mid 1960s Garfield (1982) assembled a list of over 60 forms of psychotherapy. A decade later the Research Task Force of the National Institute of Mental Health estimated 125 different forms (1975) and in 1986 Kazdin made reference to over 400 variants of psychotherapy. A large number of these have been investigated empirically.

Although a comprehensive review of non-psychodynamic psychotherapy research is beyond the scope of this thesis, there are a few areas that are relevant to our research and not adequately covered in the preceding review of psychoanalytic outcome studies. First, it is useful to review major themes within the field and discuss the perspective of some leading researchers on what psychotherapy research has achieved and what it must strive to become. Second we review three theoretical questions relevant to the psychotherapy process and outcome literature that play important roles in the design of the study described in this thesis: (1) what is meant by "structural change" and how does it differ from symptomatic and diagnostic change? (2) what is the "process" of

psychotherapy and psychoanalysis?, and (3) what is “attachment theory” and how does it relate to psychopathology and the effect of psychotherapy? Finally, we will discuss practical methodological issues relevant to this study, including principles of research design, a core battery, and methods for assessing change.

2.2 Major themes within psychotherapy research

2.2.1 Eysenck's challenge

No question is more fundamental to the field of psychotherapy research, than whether psychotherapy (in any of its forms), is capable of improving the lives of its recipients (however this may be measured). In 1952, British psychologist Hans Eysenck challenged this simplest of claims, by reviewing 19 studies in the literature, involving over 7,000 patients, and concluding that available evidence did not support the claim that patients in psychotherapy improved more than non-treatment controls (Eysenck, 1952). This finding was based in large part on the assertion that “about two-thirds of severe psychoneurotics show recovery or considerable improvement without the benefit of systematic psychotherapy [or any alternative treatment], after a lapse of two years from the time that their disorder is notified, or they are hospitalized” (Eysenck, 1952, p. 711). In order for psychotherapy to demonstrate efficacy, it would have to show rates of improvement significantly higher than 67%, something almost no study has been able to do.

Eysenck's methodology has been successfully challenged by numerous psychotherapy researchers (Garfield & Bergin, 1994). Kiesler (1966) was particularly effective in dismissing Eysenck's claims by showing that the high rate of spontaneous remission that he estimated is based only on two ambiguous studies. However, Eysenck and his 1952 paper have long served as concrete opponents for psychotherapy researchers, motivating research that would demonstrate that psychotherapy is effective.

In the past 50 years, researchers moved from arguing that methodology was still insufficient to support Eysenck's claims, to the more general finding that recipients of psychotherapy do improve more than untreated controls.

2.2.2 *The dodo bird verdict*

Almost as central to psychotherapy research as Eysenck's challenge of whether therapy works at all, has been the question of whether research could show *which* forms of psychotherapy were the most effective. Despite repeated efforts to show such differential effects, the majority of studies have failed to find statistically significant differences between the efficacies of a variety of therapies, and major figures in the field have concluded that such differences may not exist. In 1936, Rosenzweig (1936) proposed that different forms of psychotherapy work equally well because of their common factors, and quoted the dodo bird from Lewis Carroll's *Alice in Wonderland* to make his point, "At last the Dodo said, 'Everybody has won and all must have prizes.'" In 1975, Luborsky (1975) compiled comparative studies from the previous 25 years and argued that "the dodo bird verdict" was confirmed. Later meta-analyses have led to similar conclusions (Lambert & Hill, 1994; Smith, Glass, & Miller, 1980; Wampold et al., 1997), and in a recent review, Luborsky (2001) has held by his earlier assertion that differential treatment effects are not supported by the literature.

Interpretation of the dodo bird verdict as evidence that psychotherapeutic treatments have equal efficacy is, however, far from unanimous. In his 1975 meta-analysis, Luborsky points out that equivalence of the "amount of improvement," is not the same as equivalence of the "quality of improvement," and with regard to assessments of psychoanalysis, he notes that meta-analyses have included far fewer dynamic and long-term treatments (2001; 1975). Kiesler (1966) points out that the findings may also result from major unsolved methodological issues, and Beutler (1995) labels this misinterpretation of a negative research finding "the myth of therapy outcome

uniformity.” Fonagy (2001e) dismisses the dodo bird verdict as most likely due to methodologic problems of psychotherapy research and criticises its effect on the field as “unhelpful and encouraging complacency.” Most convincingly, he points to outcome studies in which the dodo bird verdict is not confirmed (Fonagy, Target, Cottrell, Phillips, & Kurtz, 2002; Roth & Fonagy, 1996), and predicts that with improved methodology, these studies will become more common.

2.2.3 Kiesler's myths of psychotherapy research

In 1966, Kiesler published a paper attempting to pave the way for more order in the field of psychotherapy research by outlining and debunking commonly held assumptions that he felt to be myths. Despite the fact that this paper was written more than 35 years ago, its lessons have continued to be as relevant today as they were when the paper was written. Kiesler begins by addressing what he believes is the mistaken assumption that patients and therapists are sufficiently uniform in their behaviours in and response to therapy that they may be combined haphazardly into groups for large scale studies (myth #1 = patient and therapist uniformity). The problem of patient non-uniformity is exacerbated by the lack of rigorous patient selection criteria in most studies (1966) and is dealt with only by stratifying or matching patients along any variables believed to influence outcome.

The non-uniformity of therapists and therapeutic techniques, even within a supposedly uniform treatment condition, is equally if not more harmful to the results of psychotherapy outcome studies (Kiesler, 1966). Kiesler suggests developing methods for identifying and measuring therapist variables relevant to outcome (therapist personality, technique, relationship, role expectancies, and the like) and then building detected differences into study design through stratification or matching. The problem is further compounded by the fact that different schools of psychotherapy were often developed with dissimilar patient populations.

Kiesler is methodical in refuting Eysenck's claim that psychopathology undergoes spontaneous remission at rates of 60 to 70% in 2 years (myth #2 = spontaneous remission) (Kiesler, 1966). Citing Rosenzweig (1954), he argues that the two studies that Eysenck uses to argue for a high rate of spontaneous remission in "psychoneurotic" patients (Landis, 1937 and Denker, 1947), each fail to meet essential criteria for psychotherapy control groups. The populations in these studies were very diverse in terms of psychopathology, the patients were not treatment-free as they were regularly seen and treated by their own physicians "with sedative, tonics, and reassurance" (Eysenck, 1952), and the criteria for improvement in these studies were not directly comparable to those used in psychotherapy studies.

Perhaps Kiesler's most scathing criticism of the psychotherapy research program deals with the lack of adequate theoretical formulations for building models of how measured variables interact and predict patient outcome (myth #3 = adequacy of present theories) (Kiesler, 1966). He accuses all three prevailing theories of his day (Freudian psychoanalysis, Rogerian theories of relationships, and behaviourism) of being deficient in providing predictions for how the network of independent, dependent, and confounding variables influence one another.

In addition to the three overarching weaknesses that Kiesler pointed to in psychotherapy research, he mentions a few other misconceptions that are troublesome to the field (Kiesler, 1966). He raises the concern that process and outcome studies are treated as different types of studies, and encourages one to always accompany the other. He questions the usefulness of current diagnostic schemes for the purposes of relating type of psychopathology to treatment outcome, and proposes, instead, that the patient's behaviour *within the psychotherapy sessions* be used as the most reliable source for diagnostic information predictive of outcome. Finally, he chides psychotherapy researchers for perpetually planning and waiting for "the definitive study" that will demonstrate efficacy

and its relationship with patient and therapist variables instead of conducting smaller, less ambitious studies for answering fewer questions.

In 1995, Kiesler revisited his original article, noting that an important outcome of his article, and several other pieces around the same time, had been a straightforward summary of the question to which psychotherapy research should devote itself: “*What* treatment by *whom*, is most effective for *this* individual with *that* specific problem, and under *which* set of circumstances” (Kiesler, 1995; Paul, 1967). He proposes a “biopsychosocial grid model” for planning future studies in which a matrix of patient, therapist, intervention, and course-of-time outcome variables is designed according to a standard set of measures and criteria, and each study aims to fill in the results of cells within this grid. Individual studies are free to either concentrate on a few number of cells by controlling variables carefully, cover many cells by introducing a factorial design, or even collapse along certain dimensions to provide more general information about a particular column or row. Combined with process research and better conceptual models, he hopes that this methodology would lead to gradual progress in the field.

2.2.4 PTO congruence and ATI research

Kiesler’s “grid model” and call for better theories foresaw another major theme within psychotherapy research: the goal of finding a conceptual model for psychotherapy that unites patient, therapy, and outcome variables in a single framework. Strupp, Schacht, and Henry (1988) named this the principle of Problem-Treatment-Outcome (PTO) Congruence and proposed that the intelligibility of psychotherapy research depends on weaving together measures from each of these domains in our models. Borrowing from Cronbach’s educational work, other authors have introduced the notion of Aptitude-Treatment Interaction (ATI) research for studying which patients do the best with which types of therapy (Dance & Neufeld, 1988; Kiesler, 1995; Snow, 1991). It is important to note that neither PTO congruence nor ATI imply that the goal of

research is to successfully match therapeutic modalities and appropriate patient populations. Kiesler points out (1995) that this may be impossible, and that it is wiser to use data about patients and therapies in concert with process measures to develop theories for mechanisms of change.

2.2.5 Abuse of the Drug Metaphor

Like many psychotherapy researchers, Stiles and Shapiro (1989) express a dissatisfaction with the current state of the field and, in 1989, put forth a theory for understanding how researchers have been led astray by improper assumptions. “After four decades of effort, and despite great advances in methodological and conceptual sophistication, we do not see commensurate substantive progress – in the form of strong, replicated findings demonstrating the sorts of process-outcome causal links that most treatment models propose” (Stiles & Shapiro, 1989, p. 521). They argue that influenced by the worlds of medicine and psychopharmacology, researchers have implicitly or explicitly endorsed a “drug model” for psychotherapy in which a psychotherapeutic treatment is analogous to a pill, its strength to the drug’s dosage, and its integrity to a drug’s purity. Manualization, according to this characterization, is an attempt to specify the ingredients of a particular therapy, as you would analyse and purify the content of a pill (Luborsky, 1984; Orlinsky, Grawe, & Parks, 1994).

Stiles and Shapiro (1989; 1994) critique seven popular assumptions they attribute to the drug metaphor, and point to the manifestations of their research in mainstream reviews of psychotherapy research (Luborsky, 1984; Orlinsky et al., 1994) as well as in their own work (Shapiro & Shapiro, 1982). These assumptions are: (1) process and outcome are distinct phenomena and can be measured as such; (2) process component names – including interpretation, confrontation, reflection, self-disclosure, challenging assumptions, focusing on affect, efforts to give support, empathy, warmth, and genuineness – signify pure ingredients; (3) process instruments measure active

ingredients of treatment; (4) active ingredients are contained in the therapist's behaviour, as opposed to within the unique therapist-client interaction, (5) the dose-response curve is ascending and linear; (6) the best way to demonstrate a procedure's efficacy is by controlled clinical trial; and (7) a process component's efficacy is shown by its correlation with outcome.

While items 1 through 5 on Stiles and Shapiro's (1989) list identify problems that can be overcome with more sophisticated models and enhanced collection of process and follow-up data, the final two items challenge the study design of many past and current efforts in the field. They argue that given the complexity of process elements, it is unrealistic to conduct clinical trials in which groups differ on a single process component or even on a host of "ingredients" in a factorial design. Meanwhile correlations between process and outcome measures are vulnerable to confounding factors and do not demonstrate causality. Absence of significant correlations is as likely to stem from methodological problems as it is from lack of an effect and can rarely be used to argue against the importance of a treatment component (1989). Other psychotherapy researchers caution against dismissing too easily the useful methods of multivariate design and correlational analysis, and suggest instead supplementing these methods with growth curve and power analyses (Sechrest, 1994; Silberschatz, 1994). Stiles dismisses these attempts, though, as avoiding the fundamental problem of trying to understand psychotherapy without a well integrated model of process and outcome (Stiles, 1994).

2.2.6 Empirically supported therapies

The most compelling recent framework for a discussion of psychotherapy research involves the debate over a movement to use research to establish a set of "empirically supported treatments" (APA Task Force on Psychological Intervention Guidelines, 1995; Chambless et al., 1996; Kendall, 1998; Safran & Aron, 2001; Westen & Morrison, 2001). In the early 1990's the Clinical Psychology Division of the American Psychological

Association appointed a task force to create a list of empirically supported treatments (ESTs) together with a set of criteria and a procedure for identifying such treatments (1995). They subsequently published the procedures they chose together with a list of 22 “well-established” and 26 “probably efficacious” ESTs (1995; Chambless et al., 1996). The criteria for selecting well established treatments, *all* of which must be met for a treatment to qualify are: (1) each had to be demonstrated as better than a pill placebo, psychological placebo, or other treatment, or as good as an established treatment in (a) a good group-design study with adequate statistical power or (b) a large series of single case design experiments; (2) experiments must have been conducted with treatment manuals; (3) characteristics of patient samples had to have been clearly specified; and (4) the evidence in favor of a treatment’s efficacy had to originate from at least two different investigators or investigatory teams. The criteria for probably efficacious treatments, *any* of which would be sufficient for a treatment to qualify, are: (1) two experiments showing the treatment more effective than a waiting-list control group; (2) a treatment meeting criteria 1a, 2 and 3 of a well-established treatment, but has not been replicated by separate investigators, and (3) a small series ($n=3$) of single case design experiments meeting criteria 2 and 3 of a well-established treatment.

As might be expected, the action of this group caused a controversy in the field, surrounding both the means by which they selected ESTs and the therapies that were selected. Much of the argument surrounds the fact that all 48 ESTs are short-term treatments, and the overwhelming majority can be categorized as cognitive-behavioural or behavioural in approach. Only one of the “well-established” ESTs (Klerman and Weissman’s interpersonal treatment of depression) is at all associated with psychoanalysis, and brief psychodynamic treatment is listed as a “probably efficacious” treatment for substance abuse and depression (Safran & Aron, 2001). If nothing else, then, the effort has served to incite debate about psychotherapy research methodology

and create a climate in which researchers realize that their results may be having real consequences for therapy selection and reimbursement.

In reviewing the ongoing efforts to select ESTs, Westen (2001) agrees that the short-term focal psychotherapy ESTs do appear to be useful for certain disorders, but warns that methodological factors may be unfairly excluding longer-term and more insight-oriented therapies. A careful analysis of these factors, suggest that existing ESTs are not clearly the treatments of choice for depression or generalised anxiety and that community studies of psychotherapy effectiveness are needed to better establish treatments of choice for these disorders. He draws attention to the distinction between efficacy studies (which assess outcome under highly controlled conditions in order to maximise internal validity) and effectiveness studies (which assess outcome as practised by therapists in the field in order to gain external validity and generalizability), and points out that the criteria so far established for ESTs emphasize efficacy designs despite their inherent weaknesses (see also Fonagy, 2001b). The more successful an efficacy study has been in eliminating heterogeneity in patients and therapies (in response to critiques such as Kiesler's) the more likely the study will be acceptable by EST criteria, yet the less likely it will reflect the real-world patients and therapy to which the clinicians and insurance companies will aim to apply them.

In keeping with the traditions of Kiesler and the drug metaphor literature, Westen describes six assumptions of the efficacy trials on which the ESTs are based (2001). He believes that these assumptions are either unproven or downright false, calling into question existing ESTs and suggesting that other forms of therapy may have been unfairly excluded from the list. These assumptions are: (1) Psychological processes are highly malleable (if short-term psychotherapies are as successful as the EST literature indicates, this must be true), (2) Most patients have one discrete problem or can be treated as if they do (underlies any study that focuses a single treatment of a single

disorder), (3) Psychological symptoms can be understood and treated in isolation from the personality of the individual who bears them, (4) Patients are able and willing to report at the beginning of treatment what is bothering them (in order for short term ESTs to work, the patient must present with a discrete symptom for which they are given treatment), (5) The elements of efficacious treatment are dissociable from one another and additive (central to the dismantling strategy outlined in the next section), (6) The elements of efficacious treatment can be spelled out in manualized form, and the interventions specified in the manual are the ones that are causally related to outcome.

Several leading psychoanalytic researchers have questioned the appropriateness of EST methodology for demonstrating the effectiveness of psychoanalysis and psychodynamic psychotherapy (Blatt, 2001; Luborsky, 2001; Safran & Aron, 2001; Strupp, 2001; Wallerstein, 2001). They point to the difficulty and inappropriateness of manualizing a dynamic therapy and the challenge of developing measures for the sort of structural, long-lasting change to which psychoanalysis aspires. In place of the randomized controlled trials favored by the EST literatures, Strupp (2001) proposes “research-informed case histories” and Blatt (2001) advocates changing the term to “systematically supported treatments” in recognition of the fact that non-empirical methods are also valid in supporting effective treatments. Luborsky (2001) further asserts that selection of ESTs is inappropriate when the literature does not demonstrate the efficacy of one form of treatment over another, and is overrun by the effects of researcher allegiance (see below). In this group, Fonagy (2001e) stands alone in asserting that manualization of psychodynamic therapies is a feasible and useful effort, that the literature does support differential effects of therapies, and that psychoanalytic researchers can rise to the occasion by performing well-designed studies to give psychodynamic psychotherapies their proper place on the list of ESTs.

2.3 Theoretical issues in psychotherapy research

2.3.1 *Theory of structural change*

It has long been an accepted wisdom of psychoanalysis that meaningful change brought about by the psychoanalytic process goes beyond the amelioration of symptoms and involves the modification of a deep-rooted entity called “character”, “personality” or “structure.” (Appelbaum, 1994; Frances, 1982; Sandler & Dreher, 1996; Sundin & Armelius, 1998; Werman, 1989). There have been many historical, theoretical, and practical justifications for this position. Freud and other early psychoanalytic clinicians noted that addressing individual symptoms by “making the unconscious conscious,” as dictated by the topographic theory, often did not lead to more permanent changes in patterns of behaviour, and were part of the motivation for development of the structural theory. More recently, the concept has been used as a way of distinguishing the far-reaching and long-lasting goals of psychoanalysis from the supposedly more superficial and short-lived achievements of psychotherapy, particularly of the supportive or cognitive-behavioural kind (Frances, 1982; Werman, 1989). Many a psychoanalytic researcher has cited the lack of an adequate measure of structural change as the chief, and occasionally the only, significant factor in preventing them from demonstrating the efficacy and superiority of psychoanalytic treatment, making such a measure a holy-grail of psychoanalytic research methodology (Fonagy et al., 2001; Malan, 1973).

For any of these purposes, it is first necessary to establish a reasonable definition and theoretical approach to psychic structure and structural change, a task that has been disproportionately neglected given how commonly the term is used (Rangell, 1989). Rapaport (1959) defined structure broadly as “configurations of a slow rate of change”, Schwartz (1981) suggested “relatively stable organization of events that is said to underlie a meaningful sequence of actions or mental phenomena”, and Kernberg (1976) proposed “stable configurations of unconscious self- and object-representations, linked by affects

and giving rise to predictable fantasies and expectations of interpersonal interactions.”

Each definition captures the relative permanence of psychic structure, and, to varying degrees, reflects an underlying theory of the mental apparatus. Aspects, though, have led to disagreements. Rangell (1989) emphasizes that structures are “clusters of functions sufficient in degree and cohesiveness to constitute psychic ‘systems’” which are not to be confused with memories, self or object representations, or introjects as sometimes used by Kernberg. Stolorow (1978) and DeWald (1972) object to the reification of structure as describing functional units of the mind, and prefer to see it as a symbolic representation or metaphor for a stable set of automatic, repetitive functions.

Far more work than can be reviewed here has been devoted to the delineation of the nature of structural change, sometimes more generally referred to as “the goals of psychoanalysis” (DeWald, 1972; Horowitz, Kernberg, & Weinschel, 1993; Ingram, 1983; Rangell, 1989; Sandler & Dreher, 1996; Shapiro, 1991; Sundin & Armelius, 1998). Summarized below, are three representative descriptions of structural change spanning a continuum from the more metapsychological and idealized (DeWald), through an intermediate stage (Rangell), to one that is more experience-near and functional (Ingram) (Blum, 1992). We begin with DeWald (1972) who sketches a mainstream classical position on structural change as a result of the working through of the transference neurosis. As transference and therapeutic regression continue, the repression barrier is lifted and core structures from infancy and early childhood appear in manifest behaviour. Working through these behaviours in the transference modifies derivatives of core elements and should be accompanied by resolution of conflict and modification of structured behavioural patterns outside the analysis. Structural change may involve the dissolution of pathological structures and establishment of new healthier ones. He further breaks down the specific (albeit metapsychological) changes one might expect from various components of the mind, based on Freud’s structural theory. Modification

of the id is reflected in the quantitative distribution of drive intensity and the intrapsychic management of discharge. Change in the superego is assessed by “the degree to which primitive and primary process introjects are replaced by secondary-process, reality-oriented, and personally developed systems of moral values” (p. 312) and improvement is reflected in reduction of neurotic guilt and correspondence of the ego ideal to realistic goals. Structural change in the ego corresponds to modification of microstructures including use of the reality principle, stability of self and identity, consciousness of the defences, age appropriate and constant object choices, appropriate drive discharge and gratification, toleration of frustration and anxiety without regression, and sublimation. All of this is seen in the use of effective and adaptive functions and the tolerance of stress without resorting to pathological responses. DeWald points out that the permanence of these changes (ultimately what distinguishes them as real structural change instead of transient psychic shifts) is based on how well-established they were during treatment, how much self-sustaining reinforcement they provide, how compatible they are with continuing environmental reinforcement, and how actively the transference was used as a motivating force.

Other theorists have chosen to describe structural change in slightly less metapsychological and more behavioural terms, though still firmly rooted in a psychoanalytic theory. Rangell (1989) lists 14 mechanisms or functions of psychic structures that are subject to modification through the psychoanalytic experience: (1) the store of conscious memories is enlarged through free association, (2) repressed memories of traumatic and unassimilable events are recovered, (3) traumatic aspects of repressed memories are made conscious, (4) the ego receives the anxiety signal less frequently, (5) fewer memories elicit caution or need for defence, (6) increase of secondary process over primary process, (7) the superego is more attuned to ego- and culture-syntonicity, (8) the id exerts less peremptory pressures for discharge, (9) improved self-representation within

the ego, (10) the repetition compulsion becomes less id-oriented and more ego-directed, (11) decisions remain steadfastly for the patient to exercise, (12) “a new analytically treated ego sits astride an altered intrapsychic process” and operates more in the conscious realm, (13) patients internalise the analytic function and anticipate altered and enhanced ego-functioning in their moment-by-moment lives, and (14) patients retain the self-analytic capacity not only for anticipated actions but also to understand affects and moods.

Ingram (1983) takes this theoretical operationalisation of structural change one step further in outlining four perspectives from which change can be identified. Alterations must be lasting and go beyond simple psychic shifts and insights (time perspective). They must correspond to changes in predominant mood, attention to practical details, and enhanced physiology, not just a change in the external environment (adaptational perspective). They lead to richer dreams, more subtle affective communication, more creative thinking and symbolization, and richer life experience, not merely more intense affect (inner aliveness perspective). Finally, structural change leads, within the therapeutic relationship, to relaxation of defences, immediacy of involvement, and resonance with the analyst’s psychic structure, not to be confused with superficial behavioural changes or countertransference effects (clinical relationship perspective).

A number of theoretical challenges have been raised to the use of concepts, such as those listed above, for describing the process of change in psychoanalysis. First, Kennedy and Moran (Fonagy, 1999b; 1991) suggest that a distinction be made between long-term aims of psychoanalysis and the methods or short-term goals by which these are reached. This suggests that clinicians, theoreticians, and ultimately researchers must be careful to distinguish what is good practice and a useful perspective in the short term, from what is expected as the end result of psychoanalysis. As suggested by Anna Freud, “It is very much like driving somewhere. Your aim is to arrive, and if instead of looking at the road,

you think how nice it will be when you arrive, you will probably have an accident” (Sandler, Kennedy, & Tyson, 1980, p. 251). Although it is often difficult to separate short and long-term aims in psychoanalysis, and theoretical writings do not always make this clear, the research literature has been quite good at drawing a distinction between measures of structural change and those of process. That tradition will be maintained here and the theory and technique of process assessment will be treated separately below.

A more recent challenge to the structural change literature has grown out of the frustration of researchers trying to develop a reliable and valid measure in concert with data that suggest high correspondence between symptomatic and structural measures. Many have pointed out that this distinction is an inherently ephemeral one for research because we inevitably use changes in observable behaviours and symptoms to infer structural changes (DeWald, 1972; Frances, 1982; Werman, 1989; Mintz, 1981). Werman (1989) also cites the evidence that structural change occurs also outside of psychoanalysis as part of his argument for a “deidealization” of this concept and emphasis on behaviour changes to benefit research and ultimately the results of psychoanalysis with patients. What is lost in this argument, is that conventional measures of symptoms and behaviour, lacking a grounding in a theory of change or structural hypothesis, usually rely on such a broad range of symptoms or imperfect data collection measures (such as self-report questionnaires), as to render them insensitive to the subtle and specific evidence of change in an individual patient. Adaptive and maladaptive patterns of behaviour in relationships and work, particularly when the subject has limited awareness of the problems and has activated unconscious defence mechanisms, are not easily reflected in general symptom measures. However, a psychoanalytically informed measure could specifically seek out the information needed to discern such patterns and through directed study of manifest behaviours make conclusions about more permanent and unconscious structures. Structural change, seen this way, is not attempting to avoid the

importance of behaviour, but rather focuses on a specific set of behaviours that is believed to signify more deeply entrenched, wide reaching, and permanent mechanisms.

It is important to recognize that while psychoanalysis may have a privileged position with regards to the theory and study of structural change, it by no means has a monopoly in bringing it about. Analysts and researchers have described convincing cases meeting all the necessary criteria for structural change as a consequence of incomplete analyses (Schlessinger & Robbins, 1974), supportive therapies (Appelbaum, 1994; Wallerstein, 1986), and people not in psychotherapy at all who are responding to beneficial life events (Appelbaum, 1994; DeWald, 1972; Frances, 1982; Werman, 1989). For psychoanalytic research, it is particularly important to note that, theoretically, the means by which structural change is brought about in the normal maturation of children have much in common with the techniques of supportive therapy. These include (1) promoting the maintenance of optimal level of arousal for learning, (2) fostering a sense of self and sense of the other, (3) encouraging mature interpersonal relationship through modeling, (4) fostering anticipatory anxiety and a sense of control over events of psychotherapy, and (5) strengthening defences that are needed at any give moment or phase of psychotherapy (Appelbaum, 1994). The differential effect of psychoanalysis and supportive therapy is probably greatest when one or more of these functions is blocked by entrenched pathological mechanisms, such as maladaptive defences or relationship patterns. It also seems apparent, as demonstrated by the Menninger Psychotherapy Research Project and much subsequent research (Jones, 2000), that supportive elements play an important role in much psychoanalytic work and, in the right circumstances, are essential to structural change. All of this raises a challenge for psychoanalytic research to prove not only that structural change has taken place, but that it is predicted by and related to psychoanalytic work. Frances (1982) quotes Edith Jacobson, "If life can

achieve so much, analysis should be able to do even more.” The question remains, how to prove this.

Two further arguments emerge from the structural change literature, though their truth value appears to be largely dependent on empirical validation, as described below. First, the enormous diversity of theoretical writing suggests that in order for a measure to usefully correspond to underlying mental structures, it should emerge from a coherent theoretical framework. A tool that borrows concepts from a broad range of theories is in danger of losing its connection to structure and to appear more like a generic measure of symptomatology. Furthermore, a coherent measure would serve as an honest representation of the theory on which it is based, and the results of studies that used it can be interpreted in this manner. Positive and negative results from a study using measures formed from an amalgam of theories are more difficult to interpret, and negative results point less clearly to an improved study.

Second, one of the most consistent areas of focus in the theoretical literature on structural change, that is not already captured through symptomatic measures is the area of relationship patterns and object representations. Though generic psychotherapy outcome measures may capture crude self-report estimates for how functional these relationships are, particularly in the highly pathological range, they miss the subtleties, unconscious patterns, and large diversity within an intermediate range of functioning. Models of structural change such as Ingram’s (1983) have important elements of object relations from multiple perspectives, including the adaptation, inner aliveness, and clinical relationship. It appears likely that an effective measure of structural change will make the study of relationships and their representations a high priority.

2.3.2 Theories of therapeutic process

As with structural change, the discussion of what is meant by psychoanalytic process and how it might be measured originates in the early days of psychoanalysis and

has generated a large theoretical and empirical literature. Virtually every major psychoanalytic theorist and researcher has at some point contributed to the debate and the field is littered with theories and measures that are always interesting, but never universally accepted or well-studied. A review of the theories of psychoanalytic and psychotherapeutic process is beyond the scope of this chapter, but three contemporary theories are presented to illustrate the variety and similarities in the field. Some general issues regarding measuring psychoanalytic and psychotherapeutic process will be introduced before proceeding to a review of some current process instruments and their advantages and disadvantages.

Two principal lines of thought about the mechanism of therapeutic action have been organized around the roles of (1) therapist interpretation and (2) interpersonal interaction as the central active ingredients (Jones, 1997). Early theories, beginning with Freud, emphasized the importance of the interpretation as bringing about lasting change and suggested that therapies that relied on a benevolent relationship between therapist and patient were unlikely to achieve permanent effects. Several theoretical and empirical factors, however, made some psychoanalysts move away from this classic notion and emphasize the importance of the relationship. Freud himself, while still emphasizing the verbal interpretation, acknowledged the importance of the transference relationship and referred to psychoanalysis as a “cure through love.” Winnicott and Kohut added to the understanding of how the relationship could be responsible for more than passing change, and Wallerstein (as described above) pointed out that lasting change could occur even in the context of an ostensibly supportive therapy.

Several contemporary theorists and researchers have suggested how the roles of interpretation and relationship can be seen together in a single theory of therapeutic action. Jones (1997; 2000) theorizes that “specific and repetitive interaction structures,” conscious or unconscious, verbal or nonverbal, contribute to patients’ capacity for

reflection and self-understanding and thus bring about change (1997, p. 1140).

Intersubjectivity, transference-countertransference enactments, and role responsiveness may all be seen as examples of these repetitive structures. He argues that measures of analytic process must take into consideration that these structures come in many forms and must be studied in both patient and analyst simultaneously for meaning to be derived.

Several Boston analysts, part of the “process of change” group, have similarly constructed a theory that incorporates multiple dimensions of the therapeutic process, including but not restricted to interpretation and relationship, as parts of a single mechanism. They describe “implicit relational knowing” in the interaction as raising internal object relations to a more general representational systems conception, integrating affect, fantasy, behavioural, and cognitive dimensions (Harrison, 1998; Lyons-Ruth, 1998; Modell, 1998; Nahum, 1998; Bruschiweiler-Stern, 1998; Morgan, 1998; Beebe, 1998; Stern, 1998; 1998a; Tronick, 1998b). Some small areas of this implicit knowing may become the subject of verbal articulation or transference interpretation, but there are other important examples of knowing that are not symbolically represented and may not even be dynamically unconscious, in the sense of being defensively excluded from awareness. They further articulate a theory of “now moments” and “moments of meeting” in therapy where the implicit knowing is highly engaged and brings about change (Lyons-Ruth, ; Stern).

In Fonagy’s (1999b) version of the process of change, greater emphasis is placed on object relations and developmental theories. He outlines a theory of personality pathology in which the central deficit is the inability to interpret the behaviour of others in terms of underlying mental states (i.e., impaired “reflective function” or “mentalizing capacity”). This presumably comes about, in part, because of caregivers who are unable or unwilling to reflect the child’s internal state leaving the child to be overwhelmed by

affect. The child then selectively excludes ideas associated with unmanageable feelings, including an appropriate understanding of the minds of others. Analytic therapy approaches this deficit in three phases. First, the therapist builds an alliance by permitting externalisations and waiting for moments when the patient is feeling secure. This may be accompanied by an increased sense of safety and initial symptom reduction. Second, in the moments of security, the therapist focuses on an understanding of mental states as they arise in both the patient and therapist. These interpretations or clarifications are often accompanied by a heightening of conflict and return of symptoms. Third, through the consistent and coherent elaboration of mental states, a reorganization or restructuring of the patient's representational system comes about, enriching the patient's object representations and reflective capacity. If accomplished properly, this shift causes underlying changes in procedural and implicit memory and leads to long lasting symptomatic and personality improvement.

Most significantly, the theories of Jones, Fonagy, and the Boston group have in common that they place equal emphasis on verbal and nonverbal or explicit and implicit mechanisms of the therapeutic process. A measure that successfully captures the interaction that they describe would have to capture the multiple levels of patient and therapist behaviour and identify the intersubjective themes that arise, rather than relying on simplistic classification of supportive and interpretive interventions. Jones ties his theory closely to a measure of psychotherapy process, described below, known as the Psychotherapy Q-sort, in which a third party reads a transcript or watches a video tape of a session and then rates the relative importance of a series of statements about therapist, patient, and joint behaviours (2000). Fonagy cites a large corpus of his own research in attachment, developmental psychopathology, and psychoanalytic outcome as part of his theory, but has not yet specifically studied the representational shifts or phases of therapy that he describes. His measure of reflective function (1997), outlined briefly above under

measures of structural change and in more detail below under attachment, suggests itself as a way to monitor this capacity continuously during therapy, and it is likely that it will be applied this way in the future. A measure to study the way in which therapist and patient behaviours within a session bring about the change in reflective function has not yet been proposed. The Boston process of change group, though including at least three accomplished researchers (Lyons-Ruth, Tronick, and Stern) has made the least progress in suggesting an objective empirical measure and still confines most of its work to the narrative description of therapy transcripts.

A good deal has been written in the psychoanalytic literature about the problems of objectively measuring therapeutic process (APsA, 1974; Beres, 1968; 1968a; Engel, 1968b; Fonagy, 2001b; Kanzer, 1968; Schlesinger, 1974; Wallerstein, 1968; 1971; E. R. Zetzel, 1968). Schlesinger (1974) begins his enumeration of concerns with the nature of the “two-part analytic situation,” the necessity for confidentiality, and “impossibility of understanding what goes on it by anyone who has not experienced it himself.” Both he and the American Psychoanalytic 1974 Conference on Psychoanalytic Education and Research (1974) treat as self-evident the idea that electronic audio and video recording of analytic sessions be considered a major advance in research potential, yet recognize the practical difficulties of making this a reality. The most obvious of these is that recording would intrude into the analytic situation, at best altering the analytic process in some unknowable way, and at worse rendering it no longer “psychoanalytic.” They conclude, though, that whatever the resulting situation, it is still of great relevance to psychoanalysis and is our best chance at an objective window into the analytic process. In terms of the confidentiality implications, the APsA report stresses that both patient and therapist deserve adequate protection in exchange for exposing themselves in this way. Schlesinger (1974) points out that the emphasis by some on the risk to the patient, may conceal an

even greater concern with the a loss of self-esteem on the part of the analyst in listening to the shortcomings of their own interventions.

Recording alone, of course, does little to address a number of other concerns in studying the psychoanalytic process and generates a few new ones. Audiotapes and even videotapes cannot capture all the subtleties of an interaction, particularly the emotional states and intrapsychic phenomena of therapist and patient that are presumably central to a session. While there are ways to begin to address some of these problems, such as using emotional coding schemes or even psychophysiological recording, to gather some of the desired information, information is always lost. As with all research, Schlesinger states, simplifying assumptions must be made, and the researcher's hope and assumption is that valuable information can still be gleaned using a necessarily imperfect model (1974). If, at least, the researcher has a clear idea of what he intends to study and finds ways for observers to agree, within their definition, on what has occurred, progress can be made. The next problem is that the volume of data generated by recording of entire analytic sessions is enormous, and researchers must find ways to meaningfully reduce these data into usable form. A number of methods have been suggested, ranging from confining the study to single sessions (Jones, 1993b), to using computers to linguistically process many sessions (Mergenthaler & Kächele, 1996), to using the therapists themselves or trained raters to condense the themes of sessions into a usable format.

Based on these challenges, a number of themes will be highlighted in the process measures considered for use in this study. First, the measure's underlying theory of change or lack thereof is important to understand what processes it can be used to investigate. Second, the mode of data collection and how it deals with the problem of data volume will be discussed. Relevant to both these issues is the identity of the informant (patient vs. therapist vs. observer) and the advantages and disadvantages of different data sources.

2.3.3 Basic attachment theory and research

In 1969 the British child psychiatrist and psychoanalyst John Bowlby proposed that a child's regulation of proximity to protective caregiving figures may be best understood through the model of a biological and instinctual attachment behavioural system (Bowlby, 1969). This system is one of several ethological systems (others being feeding and reproduction) necessary for survival in humans and primates. The attachment system controls the infant's continuous monitoring of the accessibility of caregiving "attachment" figures (usually, but not necessarily, biological relatives) and the infant's retreating to the protection of these figures in times of alarm.

The attachment field has been described as developing in three principle phases (Main, 1996): (1) description of the attachment system, 1969-1977, (2) investigation of individual differences in the attachment system in infants, 1978-1984, and (3) study of representational processes in attachment across the life-span, with a focus on developmental correlates of attachment related measures, 1985-present. In the first phase, Bowlby drew together evolutionary theory, observation of care-seeking behaviours in nonhuman primates, psychoanalytic theories of development, and, most importantly, observations of children during separation and reunion from their caregivers, into a coherent theory of an attachment system. This theory was first laid out in presentations to the British Psychoanalytical Society as early as 1958, to the strenuous objections of many psychoanalysts. The theory was immediately applied by Mary Ainsworth, a member of Bowlby's research team, in explaining her systematic observations of babies in Uganda (1953-1955, published in 1967) and then in Baltimore (1963-1978). In both environments Ainsworth observed distinct patterns followed by different children in dealing with separation and reunion with their mothers, and the correspondence of these patterns with maternal behaviour (Bretherton, 1995).

In 1978, Ainsworth published the results of her observational studies and outlined three organized patterns of behaviour in 12-month-olds who underwent a carefully outlined laboratory procedure involving two brief separations and two reunions with a parent (the “Strange Situation”) (Ainsworth, Blehar, Waters, & Wall, 1978). Infants identified as secure (“B”) showed signs of missing the parent on first separation, often cried during the second separation, and greeted the parent actively on reunion, seeking to be held and maintaining contact, until eventually settling and returning to play. Infants identified as insecure-avoidant (“A”) did not cry on separation, attended to toys and the environment throughout the procedure, actively avoided and ignored the parent on reunion, and were predominantly unemotional and did not express anger. Finally, infants identified as insecure-resistant-ambivalent (“C”) were preoccupied with the parent throughout the procedure, often seemed actively angry, alternately sought out and resisted the parent, and failed to settle down after reunion, continuing to focus on the parent and cry (Main, 1996).

Since 1978 dozens of studies have validated the existence of the attachment patterns as described by Ainsworth in a variety of cultures and settings. Researchers have found associations between attachment patterns and environmental determinants, principally parental caregiving and maltreatment, indicating that a significant part of an infant’s classification is determined by parental behaviour. This is evidenced in (a) the relation between attachment to a given parent and the history of interaction with that parent, (b) the prebirth predictability of Strange Situation response to a particular parent from that parent’s Adult Attachment Interview (AAI, see below), (c) the fact that changes in attachment to mother appear to be coordinated with changes in her life circumstances, (d) the predominance of security of attachment to the mother in impaired or ill, as well as healthy, babies, and (e) the independence of Strange Situation classification to mother and father (Carlson & Sroufe, 1995; van IJzendoorn, 1995a).

The study of maltreated infants led to the observation that many such infants are not classifiable according to Ainsworth's three categories. These infants show an array of anomalous or conflicted behaviours in the parent's presence, such as rocking on hands and knees with face averted after an abortive approach, freezing all movement with a trancelike expression, or rising to greet the parent and then falling prone. On the basis of these findings, Main and Solomon (1990) defined a fourth attachment category, designated insecure-disorganized-disoriented ("D").

A great deal of research has also been devoted to studying behavioural and psychopathological correlates of attachment classification in infants at the time of their classification and later as they mature. Infants classified as secure with their mothers mature into children with greater ego resilience and social and exploratory competence than children who were insecure infants. Security with fathers also predicts a favorable outcome as children (Carlson & Sroufe, 1995). Infants classified as avoidant are more likely to develop depression, affective distancing, and compulsive behaviour than secure infants, while infants classified as resistant-ambivalent are more likely to develop impulsive behaviour, agitation, anger, and suspicion (Crittenden, 1995). Disorganized infants develop the most severe problems of all four groups, showing depression, disruptive-aggressive behaviour, and dissociative disorders (Lyons-Ruth, 1996; Main, 1996). Disorganized infants also show the greatest levels of stress during the Strange Situation itself, as evidenced by increased cortisol levels (Spangler & Grossmann, 1993).

In the course of studying precursors and sequelae of infant attachment patterns, the focus of research shifted towards measuring internalised representations of relationships and relating these to observational evidence of relationship patterns, such as the Strange Situation. Researchers first found that drawings and narratives about relationships in middle childhood were associated with patterns of attachment as measured years before (Grossmann, 1995). The most significant step in this direction,

though, has been Mary Main's development of the AAI, a structured, semi-clinical interview focusing upon early attachment experiences and their effects. Subjects are asked for five adjectives to describe their relationship with each parent during childhood, memories to support each of these adjectives, whether they felt closer to one parent and why, whether they felt rejected during childhood, whether parents had been threatening, why they feel parents behaved as they did, how these experiences have affected their personality, and to describe any major loss or traumatic experiences (George, Kaplan, & Main, 1996). Completion of the AAI typically takes between 45 and 90 minutes and it is administered by an interviewer trained in the AAI, but not necessarily in psychology or clinical work. The interview has two main aims: (1) to "activate" the attachment system and "surprise the unconscious" by asking detailed questions about relationships in early life which the subject is unlikely to have answered before, and (2) to focus on details and review the same material from different perspectives so as to give the subject ample opportunities to contradict, or fail to support, earlier or succeeding statements. Successful completion of these goals is theorized to optimally reveal a subject's non-conscious, procedural patterns for representing relationships and managing the behaviour and emotions that accompany them (Crittenden, 1994).

The tape-recorded AAI is transcribed verbatim, with great attention to capturing pauses, incomplete sentences and words, and speech errors, and then subjected to a detailed manualized analysis (Main & Goldwyn, 1985/1994). The analysis proceeds in three stages: in the first and second stages the judge uses a set of nine-point rating scales in an attempt to ascertain (1) the speaker's probable childhood experiences with each parent (e.g., loving, neglecting, involving/preoccupying), and (2) the subject's current state of mind with respect to these experiences as revealed in discourse usages (e.g., overall coherence of transcript, vague discourse, lapses in monitoring of reasoning or discourse during the attempted discussion of traumatic events). In the third stage, the

judge considers the transcript as a whole, as well as individual experience and state of mind scores, in classifying the transcript as exhibiting one of four main attachment patterns, each with several subtypes.

The AAI classification patterns were developed theoretically and empirically to be related to classifications on the Strange Situation received by infants of the mothers to whom the AAI was administered. The patterns are derived not from apparent life history, but rather from discourse usage involved in the presentation of that history as identified by the linguistic philosopher Grice (1975). Grice suggested that rational, coherent, and collaborative discourse is most likely to be achieved when speakers adhere to four maxims: quality (“be truthful, and have evidence for what you say”), quantity (“be succinct, and yet complete”), relation (“be relevant to the topic as presented”), and manner (“be clear and orderly”). In coding of an AAI, particular attention is paid to violations of these maxims, which correspond, if severe enough, to distinct patterns of attachment insecurity.

A speaker’s state of mind with respect to attachment is classified as secure-autonomous (“F”) when the presentation and evaluation of experiences is internally consistent, and responses are clear, relevant, and reasonably succinct, regardless of whether the experiences described are favorable or unfavorable. Speakers maintain coherent, collaborative discourse, without major violations of Grice’s maxims, and convey their valuing of attachment, while remaining objective regarding particular experiences or relationships. In low-risk populations, a secure classification is found in a majority (59% in the combined meta-analytic sample of van IJzendoorn, 1995a, unresolved and cannot classify not considered, N=870) of adults, but is less common in psychiatrically distressed samples or mothers of clinically distressed children. The secure prototype was formulated by reading transcripts of interviews with mothers whose children had been classified as secure on the Strange Situation, and numerous studies

have shown a strong association between security in children and their parents, even when the parents' AAs was conducted before the child was born (van IJzendoorn, 1995a).

Secure transcripts are further subdivided into eight subgroups, ranging between tendencies toward two opposing styles of insecurity. Subjects receive an F1 classification if there is some setting aside of attachment, usually because of a difficult childhood – F1a if there has been re-evaluation and re-direction of personal life as the successor to a harsh childhood, and F1b if there is limited involvement with attachment, often as a natural consequence of a background featuring little for or attention to attachment relationships. Subjects are classified as F2 if they are somewhat dismissing or restricting of attachment, which is ultimately belied by affection, compassion, humour, forgiveness, or similar evidence. An F3 classification is the prototype of F, showing the strongest overall coherence of the transcript – F3a for individuals with largely supportive families and childhoods, and F3b for subjects with difficult experiences during childhood who are presently exceptionally thoughtful (sometimes called, “earned security”). F4 subjects show strong valuing of relationships, with some accompanying preoccupation with attachment figures, with separations or with past trauma – F4a if their childhoods were largely supportive, but accompanied by some difficult parental figure or experience, and F4b if their childhoods included traumatic experiences such as loss or abuse. Finally, subjects receive an F5 classification if they are still somewhat resentful and conflicted about childhood attachment relationships, but accept their own continued involvement and are coherent in describing it.

A speaker's state of mind is classified as dismissing (“Ds”) when Grice's maxim of quality is violated, in that one or both parents are described in uniformly positive, normalizing, and/or idealizing terms, without supporting evidence, or sometimes with active contradiction. Dismissing subjects often claim to place little value on attachment

relationships, and deny that negative experiences have had any effect on them. They also insist on lack of memory and give unusually short and uninformative answers to the interviewer's questions, violating Grice's maxim of quantity. A dismissing adult classification in a parent has been associated in numerous studies to infant avoidance in the Strange Situation (van IJzendoorn, 1995a). Incidence of a dismissing classification in low-risk samples has been found to be approximately 23% (N=870, combined meta-analytic sample of van IJzendoorn, 1995a, unresolved and cannot classify not considered).

Dismissing subjects are divided into four subgroups. Subjects are classified Ds1 if they explicitly block discourse through an absence of memory for childhood or are strongly idealizing of at least one parent. Subjects are assigned to Ds2 if they implicitly devalue attachment in their interview and describe parents with cool devaluation and derogation. Ds3 subjects are noticeably restricted in feeling and describe rejection and lack of closeness to their parents without being fully dismissing of attachment experiences. Ds4 is an empirically derived subgroup created for parents who are unusually fearful of loss of a child; they were found to have insecure-avoidant infants, even in the absence of other signs of a dismissing attachment style.

An AAI transcript is classified as preoccupied-entangled ("E") if the speaker exhibits a confused, angry, or passive preoccupation with attachment figures and is markedly noncollaborative. Entangled subjects violate Grice's maxim of manner by using excessive psychological jargon, nonsense words, or child-like speech; they violate the maxim of relevance by sometimes answering questions clearly about the past with answers about current situations and feelings, and they violate the maxim of quantity by giving long, rambling answers. Entangled classification in a parent is associated with insecure-anxious-ambivalent behaviour of the infant in the Strange Situation (van IJzendoorn, 1995a). Incidence of an entangled classification in a low risk sample has been

found to be approximately 18% (N=870, combined meta-analytic sample of van IJzendoorn, 1995a, unresolved and cannot classify not considered).

Entangled subjects are divided into four subgroups. Transcripts classified as E1 are identified by widespread passivity of thought processes regarding an ill-defined experience of childhood. Speech becomes confused, vague, or incoherent, and the reader is left with a general feeling of negative experience, without understanding the details. Subjects classified as E2 have high ratings for current, involved anger towards one or both parents, often with marked violation of the maxim of quantity. The E3 subgroup is rare in low-risk samples, but two studies have found it in a majority of subjects with DSM-III-R diagnosed borderline personality disorder (Fonagy et al., 1996; Patrick, Hobson, Castle, Howard, & Maugan, 1994). E3 subjects are fearfully preoccupied by traumatic events, usually physical or sexual abuse, traumatic loss, or a parent's psychosis. Subjects receive an E3a classification if they are confused, fearful and overwhelmed by traumatic or frightening experiences; subjects receive an E3b classification in those rare cases where they have distressing loss of memory in apparent relation to traumatic experiences.

The fourth major AAI classification is used when no single state of mind with respect to attachment is predominant in the transcript, and the transcript is designated as "cannot classify" ("CC"). This classification was introduced with the original coding manual but not emphasized until research in the early 1990's found that it was more common in clinically distressed samples, abusive and sexually abused individuals, and psychiatrically distressed criminals (Hesse, 1996). Currently 7 to 10% of transcripts in low-risk samples are designated as cannot classify. Cannot classify is typically assigned in two situations: (1) the subject changes category in mid-interview in a shocking manner, most readily obvious when the change is from entangled to dismissing or vice versa, or (2) the subject seems to be in completely differing states of mind in describing different

people. Cannot classify is *not* used to describe confusing transcripts that do not meet either of these criteria, such as that of a subject whose state of mind does not seem congruent with the experiences they described, a subject who is difficult to classify because they seem to be right on the edge between two classifications, a subject with minor elements that do not fit in with an over-riding pattern, or in the case of an interview that was improperly conducted, interrupted, or mistranscribed. If a transcript is designated cannot classify, the judge must also select one or more alternative classifications even though they do not apply well.

In addition to assignment to one of the four major categories (F/Ds/E/CC), AAI transcripts are assessed for the presence of lapses in the monitoring of reasoning or discourse during discussion of potentially traumatic events, including losses, abuse, and other trauma. Typical markers of these lapses in monitoring are errors in describing a person as having died at different times at different points during the interview, abruptly shifting to eulogistic speech when describing an attachment figure who has died, discussion of a dead person as if he/she were still alive, and report of an extreme behavioural reaction in response to a loss or trauma. Individual traumatic events are rated on a nine-point scale as to whether the subject shows lack of resolution with respect to the event, and a score above 5 on any of these events results in an unresolved (“U”) classification. In a low-risk sample, the unresolved classification is found in approximately 18% of subjects (Bakermans-Kranenburg & van IJzendoorn, 1993). Unresolved classification in a parent has been linked to disorganized infant attachment in several studies (van IJzendoorn, 1995a).

The most important step leading to inter-rater reliability of the Main and Goldwyn system for classifying the AAI has been the establishment of Main and Hesse’s program for AAI coding institutes and certification. After attending a two week, full-time course in which several transcripts are discussed and coded in a group, raters are asked to rate

30 transcripts over the next few years. Only when they have achieved adequate reliability on 30 transcripts (80% correspondence on the 4-way classification of F/Ds/E/U with ratings determined by Main and Hesse) are they certified as officially trained AAI raters. To date, most major attachment studies, including the ones described in this chapter, use only certified AAI raters. In addition numerous studies have reported good inter-rater reliability on the three- and four-way classifications. Benoit and Parker (1994) report 74% and 83% correspondence on classification of 42 transcripts of mothers and grandmothers according to the three- and four-category systems, respectively ($\kappa = 0.61$ and 0.74) compared to 33% and 35% expected by chance alone. On transcripts from 20 Israeli young-adults, Sagi and colleagues (1994) report three-way correspondence between 90% and 100% ($\kappa = 0.82$ to 1.0). Bakermans-Kranenburg and van IJzendoorn (1993) report agreement of 81% and 75% on three- and four-way categories, respectively, ($\kappa = 0.72$ and 0.66) on a sample of 16 mothers.

Three major studies have measured test-retest reliability and discriminant validity of the AAI, with most results indicating that the interview is a stable test of current state of mind with respect to attachment. Bakerman-Kranenburg and van IJzendoorn (1993) repeated the AAI after two months in 83 mothers, yielding 78% and 61% agreement between the two ratings on three- and four-way classifications, respectively ($\kappa = 0.63$ and 0.43). Although clearly statistically significant, stability of some of the categories is questionable when considered individually. While 38 of 46 (83%) first time F ratings were repeated at Time 2 (with only eight subjects rated as F at Time 2 who were not rated F at Time 1), only 13 out of 17 (76%) E ratings were repeated (with two new E's) and 14 of 20 (70%) Ds ratings (with 8 new Ds's). Most worrisome, using the four-way classification, only 7 of 14 (50%) Time 1 U ratings were repeated, with eight new U ratings appearing. This may have been in part related to subjects classified as unresolved at either time because of recently experienced loss. Main and Goldwyn (1985/1994)

specifically warn against classifying subjects as U on the basis of losses experienced in the past year, and stability of U improved when this rule was followed. Benoit and Parker (1994) reported better stability results in a sample of 84 mothers, the first interview conducted one month before the baby was born, and the second conducted when the baby was 11 months old. They found 90% and 77% correspondence on the three- and four-category systems, respectively, when 55% and 38% were expected by chance alone ($\kappa = 0.79$ and 0.63). More interestingly, 97% of secure mothers had the same classification after 12 months, 87% of E mothers, 60% of Ds mothers, and 82% of U mothers (although 38% of mothers rated as U at Time 2 were not rated as U at Time 1). Finally, Sagi and colleagues (Sagi et al., 1994) report 90% correspondence ($\kappa = 0.79$) between AAI's performed three months apart on 59 Israeli male and female college students. However, the distribution of classifications at both times was somewhat different from that found in other low-risk samples: on three-way classification 69% F, 24% Ds, 4% E. Only two subjects were rated as U, despite the fact that 66% of subjects reported the loss of a significant other, similar to the rate reported in other studies with higher numbers of U transcripts.

At the time of the second interview, Bakermans-Kranenburg and van IJzendoorn (1993) asked subjects how they felt about being interviewed again, what ideas they had about the purpose of the interview, how well they were able to remember questions on the interview, whether they had discussed the interview with friends or family, whether they had reflected on the interview questions before the second interview, whether they had changed their minds about answers given in the first interview, and whether they thought they had given the same answers again. The only significant finding from these questions was that subjects who changed AAI classification reported that they had reflected more on the first interview, and, paradoxically, more often indicated that they had given the same answers. Subjects who changed classification did not differ from

other subjects in the length of the interview, frequency of reporting that something could not be remembered, or number of references to the first interview. Subjects who changed in U classification did refer to the first interview more often, but did not differ in the number of losses mentioned in each interview.

Tests of discriminant validity have shown that classification on the AAI is largely independent of the person administering the interview, age of the subject, verbal and performance IQ of the subject, quality of non-attachment related autobiographical memory, and the extent to which a person tries to present themselves in a socially desirable light (Bakermans-Kranenburg & van IJzendoorn, 1993; Sagi et al., 1994). Statistical power was found to be adequate in all such studies. Only weak associations, at best, have been shown between the AAI and content-based measures of parenting style and general personality interviews (van IJzendoorn, 1995a).

Empirical evidence has been overwhelming that there is a significant correspondence between AAI classification in adults and the classification of the attachment behaviour of their infants with them on the Strange Situation. Van IJzendoorn (1995a) report a meta-analysis of 854 parents and infants with an effect size of 1.06, corresponding to 75%, 70% and 64% correspondence between Strange Situation and two-, three-, and four-category AAI classification, respectively. On the secure-insecure split, a 50% correspondence is expected by chance, and it is believed that parental attachment accounts for 50% of the remaining variance in infant security. The correspondence is impressive considering that the two measures are so different: one a semi-structured interview involving the coding of discourse strategies and the other a structured laboratory procedure involving coding of infant behavioural responses to reunion with an attachment figure.

Some details in the AAI-Strange Situation correspondence remain to be worked out. While the AAI predicts Strange Situation using three- and four-category

classifications, it seems that the preoccupied AAI-ambivalent Strange Situation (when an unresolved category is included) is weakest. Most research has focused on mother-infant correspondence, and so far it seems that father-infant correspondence is less robust. It has also been shown that less training in applying Strange Situation classifications leads to weaker associations, suggesting that with increasing rigor in application of both AAI and Strange Situations, correspondence will improve.

Some developmental researchers (Fox, 1995) feel that although the correspondence between AAI and Strange Situation is striking, two basic tenets of attachment theory – that maternal behaviour toward her infant is a product of the manner in which she views her own relationship with her parents, and that the manner in which the adult views his or her relationship with his or her parents is the product of construction over time – remain unproven by attachment research. Fox argues that coherence of discourse about early memories, as measured by the AAI, may be a function of other, as yet untested, personality factors that may have little to do with the subjects early attachment experiences or behaviours. He believes that developmental outcome of the infant is a function less of early attachment relationships than of early individual differences, such as in temperament, that interact with a stable or changing environment. He points out that the often cited attachment finding that infants have independent attachment classifications with mother and father, independently predicted by the AAIs of those parents, does not hold up to meta-analysis or more careful data analysis of individual studies. Attachment researchers rebut these arguments by pointing out that significant associations between AAI classification and temperamental characteristics have not been found (van IJzendoorn, 1995b).

The most recent findings directly applicable to understanding the relationship between AAI and strange situation consist of the results of 16- to 20-year longitudinal studies that began with classification of infants on the Strange Situation and are now

yielding data as to the AAI classification of those same subjects. Two of three such studies have yielded a high degree of correspondence (70% to 77%) between infant and adult security vs. insecurity. In one study, the extent of this correspondence was improved when subjects who had suffered negative life events were removed (Main, 1996).

Although attachment research grew out of Bowlby's theoretical understanding of parent-infant interactions, and in particular psychoanalytic models for how this relationship affects the mental world of the child, in the last several years the attachment field has focused on accumulating new data without the necessary theories to explain them. Van IJzendoorn (1995a) raises the serious problem of a "transmission gap"; while parental AAI and infant Strange Situation are clearly related, no existing data explain the mechanism for this association. A correlated error variance is unlikely, as the two methods are so different, and measurements of neither temperament nor parental sensitivity account for the strength of the association.

Fonagy, Steele, and Steele, who were responsible for the first study relating pre-birth AAI classification to infant Strange Situation in a sample of 100 London families (Fonagy, Steele, & Steele, 1991), have suggested a synthesis of psychoanalytic and developmental theories about the development of a child's "theory of mind" to bridge the transmission gap (Fonagy et al., 1995). They explain that secure attachment is the outcome of successful mental "containment" of the infant by the mother, which includes responding to the infant emotionally, mirroring its affect, modulating its unmanageable feelings, and acknowledging its intentional stance. Insecure attachment, on the other hand, is a defensive compromise in which intimacy (avoidant/dismissive) or autonomy (resistant/preoccupied) are sacrificed for the sake of retaining physical proximity to a caregiver who is incapable of containing the infant's affect.

Fonagy and colleagues, inspired by Mary Main's proposed link between the AAI and metacognitive monitoring in the caregiver, suggest that measuring a caregiver and child's capacity for interpreting their own and other people's behaviours in terms of mental states would help fill the gap between AAI and Strange Situation. They designed a one-dimensional "reflective function" scale which is used in conjunction with the AAI to assess individual differences in this capacity and have found it to be not only highly correlated with AAI and Strange Situation, but a significant mediator of the interaction between the two (1995; Fonagy, Steele, Steele, Moran, & Higgitt, 1991; 1997).

2.4 Practical solutions to methodological issues

2.4.1 Psychotherapy research design

Numerous psychotherapy researchers have categorized the range of available research designs and discussed the advantages and disadvantages of each for answering specific questions (APA Task Force on Psychological Intervention Guidelines, 1995; Fonagy, 2001a, 2001b; Gabbard, Gunderson, & Fonagy, 2002; Goldfried, Greenberg, & Marmar, 1990; Horvath, 1988; Kazdin, 1994; Kiesler, 1966; Parloff, 1986; Paul, 1967; Richardson, 2001; Westen & Morrison, 2001). Kazdin (1994) provides a useful taxonomy along three axes: (1) evaluation strategy (treatment package, dismantling, constructive, parametric, comparative outcome, client and therapist variation, and process), (2) type of investigation (true experiment or randomized control trial (RCT), quasi-experiment, and passive-observational study), and (3) type of design (group comparison, single group, single-case, and qualitative or conceptual research). Historically, the treatment package and comparative outcome strategies, randomized control trial investigation, and group comparison design have been considered most useful. Recently, Gabbard and colleagues (2002), in their review of psychoanalytic outcome research, re-emphasized the value of RCTs in establishing empirically supported treatments, giving the following hierarchy in

descending order of usefulness: (1) RCT of study treatment versus a well established alternative, (2) RCT of study treatment versus placebo treatment (ethically questionable), (3) RCT of study treatment versus treatment as usual (specificity of treatment benefit difficult to determine), (4) RCT of study treatment versus waiting list control (still worse for determining treatment specificity), (5) prospective pre-post treatment (quasi-experimental better than passive-observational), (6) case series, (7) case report, and (8) description of clinical innovation.

However, a more general consensus from the recent literature is that research questions must be looked at individually, to determine the most practical and helpful research methodology. Despite some earlier discouraging results (i.e., dodo bird verdict), treatment package, dismantling, constructive, parametric, and comparative outcome studies are still valid means for investigating various sorts of psychotherapy as long as they select their samples and treatments carefully (avoiding the pitfalls of patient and therapist nonuniformity) and make as much use as possible of process measures. RCTs, while not practical in all circumstances and certainly not the ultimate solution for all psychotherapy research questions (Goldfried & Wolfe, 1998; Jadad, 1998; Richardson, 2001; Shadish et al., 1997; Warren & Norton, 2004), are still the most rigorous and empirically valid method for examining differences between alternative therapies. Meanwhile, quasi-experiments provide a flexibility that make them ideal for measures of effectiveness, essential to the generalizability of psychotherapy research to clinical recommendations (Grant & Sandell, 2004). Passive-observational studies will continue due to their obvious practicality, though they are the least useful in confirming hypotheses.

Future promise in psychotherapy research lies in two areas: (1) merging of rigorous RCT and quasi-experimental methodology with new process measures to create generalizable process-outcome studies that satisfy the criteria described by Westen

(2001), Fonagy (2001e), and Kiesler (1995) and (2) increasing sophistication of single case research methodology, particularly in the experimental and quasi-experimental realm, followed with direct and systematic replication (Hilliard, 1993). These principles formed the motivation for the study design described in Chapter 3.

2.4.2 Development of a core battery

The selection of appropriate measures for capturing symptomatology and change is fundamental to the effort of psychotherapy research. A quick review of the measures used over the past 50 years reveals that the situation borders on the chaotic. Froyd and Lambert (1996) reviewed 334 outcome studies whose results were published in peer review journals between 1983 and 1988, finding a total of 1430 measures, 851 of which were used only once, and 278 of which were not standardised at all. Even when studies are restricted to those dealing with a single well-defined disorder, there are as many different measures as there are studies. Ogles and colleagues (1990) found 98 measures in 106 studies of agoraphobia. A recent handbook of selected measures published by the American Psychiatric Association for clinicians includes 240 measures and runs to more than 800 pages (2000).

Needless to say, the proliferation of measures in psychotherapy research has made standardisation of results and comparison across studies extremely difficult (Fonagy, 2001c; Horowitz, Strupp, Lambert, & Elkin, 1997; Lambert & Hill, 1994). Psychotherapy researchers have long recognized this problem and advocated for a standardised “core battery” of measures, but with little result. In the early 1970’s, the Clinical Research Branch of the National Institute of Mental Health (NIMH) sponsored an Outcome Measures Project whose efforts to find a battery stemmed from the hope that “if researchers working in different setting with different treatment orientations were to use the same standard set of instruments, it would become possible to compare and integrate the results of different studies” (Waskow & Parloff, 1975, p. 3). The group published a

book with specific recommendations in 1975 (suggesting that a battery include The Psychiatric Status Schedule, the Hopkins Symptom Checklist, Target Complaints, both patient and therapist forms, the Minnesota Multiphasic Personality Inventory; and either the Katz Adjustment Scales or the Personal Adjustment and Role Skills Scales) but these seemed to be little noted by the field. Participants in the effort have since blamed this outcome on the failure to develop batteries specific to particular patient disorders and problems and the absence of an overall conceptualization about what should be included in such a battery. Twenty years later, another effort to establish a battery grew out of the Society for Psychotherapy Research (SPR) and yielded the 1997 book *Measuring Patient Changes in Mood, Anxiety, and Personality Disorders: Toward a Core Battery* (Strupp, Horowitz, & Lambert, 1997). Although this work went further in establishing a guiding structure and recommending a hierarchical system of universal, general, and specific batteries (for all disorders, general categories of disorder, and specific disorders, respectively), it still fell short of a conclusive list measures to be used in all outcome studies.

The state of a core battery of measures in psychoanalytic research is as bad, if not worse, than in psychotherapy research in general. In the *Open Door Review of Outcome Studies in Psychoanalysis*, Fonagy and colleagues (2001) list as many measures as there are studies, with little standardisation. In addition, they note the need for measures specific to psychoanalytic constructs, suggesting that their concerns need to be concerned in the development of a core battery, or that such a battery will have to be supplemented for studies of psychoanalysis.

Despite the enormous pool of available measures, researchers do agree that the set they choose satisfy a number of criteria, several of which emphasize the diversity of perspectives that a battery must represent (Elliott & Anderson, 1994; Fonagy, 2001b; Horowitz et al., 1997; Lambert & Hill, 1994; Luborsky, Diguier, DeRubeis, & Schmidt, 1997; Seidenstrücker & Baumann, 1978; Westen & Morrison, 2001). These criteria

include: (1) information should be drawn from differing sources of information (e.g., patient, therapist, patient's significant others, independent observer); (2) measures should cover a range of symptom domains (e.g., affect, cognition, behaviour); (3) information regarding change should be as specific as possible; (4) measures should draw from different domains of functioning (e.g., work, social, marital); (5) information regarding psychopathology should be captured in both categorical and continuous terms; (6) measures should be independent of particular theories and schools of psychotherapy; (7) measures must have good psychometric properties; and (8) the cost of assessment must be weighed against the usefulness of any one measure.

It has long been known that different sources of information for patient variables yield surprisingly non-unitary results (Lambert & Hill, 1994). The only satisfying approach for understanding this discordance has been to say that each informant brings with them a particular database, valuable in its own way to understanding the larger, unknowable, picture of the patient (Elliott & Anderson, 1994; Gerber, 1994; Seidenstrücker & Baumann, 1978). Strupp and colleagues (Horowitz et al., 1997) suggest viewing the discrepancies in how different sources assess change in a patient according to a tripartite model. The goal of society, as measured by information from family members and significant individuals in the patient's life, is for the patient to comply with the rules of an orderly world and fill his role appropriately. The goal of the individual patient, as measured by symptom checklists and self-report questionnaires, is to be happy. The goal of mental health professionals, as measured by therapist report, is for the patient to have a sound personality structure, as evidenced by good interactions within sessions and patient report of a happy and productive life. More typically, the necessary sources for patient information are summarized as: (1) self-report, (2) significant other, (3) therapist, (4) trained observer, and (5) instrumental (i.e., societal records or physiological instruments). Westen (2001) emphasizes the importance of behavioural sources of

information, such as whether a patient seeks treatment again after conclusion of therapy. A satisfactory assessment battery will incorporate as many of these sources of information as possible.

Subjects may also show an impressive diversity in the range of symptom domains in which they express their psychopathology and/or health (Lambert & Hill, 1994). In psychotherapy studies that do not select for a single disorder this is most pronounced, although given the comorbidity of disorders and the inadequacies of our diagnostic criteria, a range of symptoms is important to capture in any research (Westen & Morrison, 2001). To capture the range of psychopathology, symptom measures should begin by including adequate sampling of affective, cognitive, and behavioural symptoms. Researchers have frequently tried to increase the specificity of symptom measures and their sensitivity to change by focusing on the patient's primary symptom and/or the patient and therapist's specific goals for therapy (Lambert & Hill, 1994; Luborsky et al., 1997; Westen & Morrison, 2001). This, however, has led to several problems. A patient's primary symptom and goals at the start of therapy may bear little resemblance to the symptoms and goals that are seen after even a few sessions. Measures of specific symptoms or goals do not insure that these are well defined, and the units of change differ for each.

The most practical suggestions involve incorporating primary symptoms and goals, as measured by both patient and therapist, into a larger framework of assessment. Westen (2001) suggests that they be supplemented with measures of general symptomatology, adaptive functioning, personality variables, behavioural data, and objective testing of implicit associational networks. Horowitz and colleagues (1997) suggest a pyramid of batteries, beginning with a universal battery that is common to all patients, followed by a general battery that is common to patients within a particular diagnostic category (such as affective, anxiety, or personality disorders), and finally a

specific battery that targets symptoms within a narrow diagnostic category (such as panic disorder or borderline personality disorder). Such a structure removes the need for comprehensive testing of every symptom in each patient, without falling victim to the vagaries of measures in which patients and therapists identify primary symptoms and specific treatment goals.

Psychotherapy researchers have long observed that measures of functioning are essential for understanding a patient's psychopathology as well as for assessing change (Lambert & Hill, 1994). Research has shown, however, significant discordance may exist in how that patient functions in different domains. This discordance may parallel, but it not fully described by, discrepancies among informants in describing the patient's functioning. Lambert and Hill (1994) divide a patient's areas of functioning along a continuum they label "content" into intrapersonal, interpersonal, and social role performance. Others enumerate a range of interpersonal and social roles in which functioning, or the appearance of functioning, may differ significantly, including marriage, family, intimate relationships, friends, school, and work (Fonagy, 2001b; Luborsky et al., 1997; Seidenstrücker & Baumann, 1978). Influenced by recent findings in cognitive psychology that explore the overlap between implicit functioning and the unconscious mind, Westen (2001) suggests also assessing functioning outside of consciousness. Measures such as the emotional Stroop and quantification of responses to the Thematic Apperception Test (TAT) may serve to assess implicit association networks and attentional biases that are either important for predicting psychotherapy outcome or measuring change in response to therapy (Westen, Feit, & Zittel, 1999).

A number of other issues concerning measure selection have arisen through empirical research and years of experimenting with the advantages and disadvantages of different methods. Under different circumstances, psychologists have found usefulness in both categorical measures, that divide subjects into groups based on diagnosis,

“normality,” or problem areas, and continuous measures, that place them along a continuum, usually from presence to absence of a symptom or positive to negative functioning. Much has been written about the failure of diagnostic categories, though the usefulness of such measures is inescapable. Continuous measures appear more informative, but often are only useful when applied to a highly specific and operationalised concept (Lambert & Hill, 1994). Although it complicates statistical procedures and measurements of change, a comprehensive outcome battery needs to incorporate both kinds of measures (Horowitz et al., 1997).

In order for a measure to be useful in objectively comparing different types of therapy as well as in a core battery that may be applied to many different treatment conditions, it must be as independent of bias and theory as possible. This is best accomplished by eliminating therapy-specific jargon from measures and ensuring that if a measure requires a trained rater, these raters are reliable independent of their theoretical allegiance and training. Successful attainment of this goal may go a long way to undoing the troublesome treatment allegiance effect, described by Luborsky (1999) and discussed in the previous section. At the same time, one must be careful to avoid making measures so theory-free that they lack any conceptual underpinning and avoid measuring any construct that is deeper than a surface symptom. Luborsky (1997) suggests including dimensions from all major psychotherapeutic orientations in a core battery, giving each school a chance to pick a measure that shows their treatment in the best light.

A seemingly obvious, but by no means universally met, requirement for measures to be included in a standard battery is that they have good psychometric properties. Froyd and Lambert (1996) report that 278 of the 1430 measures they counted in studies lacked any standardisation at all. A good measure should have demonstrated reliability across multiple paradigms, if necessary, including test-retest within a single individual over time, inter-rater reliability for a scored measure, and split-half reliability for a

questionnaire or symptom checklist (i.e., correspondence between two arbitrarily split halves of the measure if it is designed to be a redundant measurement of the same construct). If a study cites existing reliability data performed by another research group, it must show that its methods and patient population closely parallel that of the other group, otherwise tests of reliability need to be repeated. If a methodology has multiple steps, reliability must be shown for each (e.g., in Luborsky's CCRT, reliability must be shown for selection of relationship episodes as well as for coding of individual themes in an episode). The reliability of change within a measure, is not necessarily implied, by the measure's reliability at one point in time, and needs to be examined separately (Lambert & Hill, 1994). Finally, reliability should be shown separately for patient and non-patient populations, as different patient attributes may affect the properties of the instrument. Much research has been devoted to investigating how to make a clinical measure more valid. In general, it has been found that measure reliability is associated with how concrete and operationalised a measure is, and how concrete and objectively identifiable are the patient attributes it is assessing (Gerber, 1994).

Validity, the extent to which a measure estimates or describes the dimension, phenomenon, or construct it purports to measure is also an essential attribute of any component of a core battery. This can be demonstrated by applying the measure to populations in which a clear difference is expected, such as normal and clinical samples, or by showing correlation with a previously established measure. Neither method is sufficient, however, in that patient populations are notoriously comorbid and diverse, and previous measures are rarely looking at an identical construct, otherwise there would be little rationale for a new measure. Several forms of psychometric validity have been described including content validity (whether it captures the construct it purports to), concurrent and predictive validity (whether a measure is associated with or predicts, respectively, behaviours as predicted), convergent validity (whether a measure correlated

with variables that it theoretically should), and discriminant validity (whether a measure does not correlate with variables that it theoretically should not) (Anastasi & Urbina, 1997). Many, if not all, of these should be demonstrated before a measure is accepted into a core battery.

Segal (1997) notes that given the widely accepted stress-diathesis model of psychopathology, it is not enough to use measures that only assess symptomatology at the patient's current level of stress. Such measures may be sensitive to recovery from a disorder, but do not necessarily detect vulnerability to future dysfunction (particularly if follow-up data is not collected at a sufficiently long interval after treatment has ended). He suggests using "priming" or "construct activation procedures" to reveal latent variables whose effects are discernible only under certain stressful or evocative conditions that are challenging to the individual.

The ultimate decision as to whether a measure is worth including in a standard battery is based on a careful cost-benefit assessment of how much new information the measure provides relative to how expensive, time consuming, and intrusive the measure is to carry out (Horowitz et al., 1997; Lambert & Hill, 1994; Luborsky et al., 1997). In order to minimise the costs of a measure its procedures should be clear and standardised, the measure efficient to administer, and it must be feasible to administer the measure before and after a course of therapy, and preferably, during therapy as well.

Several disadvantages of individual measures, not obvious from the above discussion, have been described in the literature. Measures completed by the primary clinician, although efficient because the clinician already knows the patient, and clinically relevant, in that the assessment is made by a person who may agree with the experimenter's conceptual framework, is fatally flawed by the confounding factor that the treatment is being performed by the same person. In addition, clinician ratings have been found to lack validity when compared with other measures (Lambert & Hill, 1994).

Measures that assess patient attributes or domains of change close to those targeted by therapy are vulnerable to artificial inflation, and are thus suspect. This emphasizes the degree to which measures must be modality independent. One danger that has been discussed in designing a battery is that even if researchers were to apply the same measure in multiple studies, there is no guarantee that the measures are being applied the same way or give the same information under different conditions. A battery should not be so strict that it “freezes the field” and does not allow for further investigation and progress in measure development (Horowitz et al., 1997).

2.4.3 Methods for assessing change

No matter how carefully we choose a study design and battery of measures that maximises our chances of answering a question about psychotherapy outcome or process, one inescapable but difficult to answer question remains: How do we go about classifying or quantifying who gets better? Historically this question was left mostly to the statisticians. If a study showed that a group of patients improved on an outcome measure in a way that was unlikely, statistically speaking, to be due to chance, they were said to be improved. However, savvy critics of psychotherapy research, particularly in the last two decades, have pointed out that statistical significance alone is not sufficient to demonstrate what they have called “clinically significant change” (Fonagy, 2001b; Imber, 1992; Lambert & Hill, 1994; Paul, 1967). Depending on the psychometric characteristics of an individual measure and its behaviour in normative samples of normal and dysfunctional individuals, a statistically significant change in individual scores or group means may or may not signify that a clinically significant change has occurred. Much attention has been devoted to developing useful metrics of clinically significant change, and some of the more common strategies, along with their advantages and disadvantages, are presented here.

Kazdin defines clinical significance as “the practical or applied value or importance of the effect of an intervention – that is, whether the intervention makes a real . . . difference in everyday life to the clients or to others with whom the clients interact.” (Kazdin, 1999, p. 332). This contrasts with “statistical reliability” which is the likelihood that an observed change in scores is not due to chance. Statistical reliability of change scores is a necessary but not sufficient condition for establishing clinically significant change. Similarly, statistical reliability of a static measure is a necessary but not sufficient condition for establishing reliability of change within that measure. One source of confusion in the literature is that some investigators use a more specific definition of clinically significant change, namely that it implies movement of scores into the normal range (Jacobson, Roberts, Berns, & McGlinchey, 1999). We will use the more general definition and discuss movement into the normal range within the context of methods for testing that hypothesis.

Before the recent growth of the clinical change literature, four principle methods were used for assessing change in psychotherapy research, each of which has been fraught with problems. First, investigators introduced one or more criteria for improvement based either on the scores of a repeated measure or on the subjective assessment of client, therapist, or an objective rater and reported the percent of subjects who met those criteria. This strategy suffered from variability in how those criteria were established, and rarely distinguished whether a change signified “recovery,” that is, eradication of the symptom and return to the attributes of a normal sample, or “improvement,” that is, a meaningful change, but not necessarily a return to normalcy (Westen & Morrison, 2001). No standard methods were used to build criteria for improvement or recovery, thus comparison of percent change from one study to the next was difficult, if not impossible.

Second, researchers calculated the difference between pre-treatment and post-treatment scores (called a “raw change score”) on a particular measure for each subject and used the analysis of variance (ANOVA) technique to look for statistical differences between group means (Cronbach & Furby, 1970). This method suffered from several statistical and interpretive failings: (1) Variability in pretreatment score confound the findings, as subjects with a worse pretreatment score have a greater opportunity to change due to floor and ceiling effects and are therefore over-represented in the result (Lambert & Hill, 1994). (2) Raw change scores are sensitive to regression to the mean, also making it more likely that subjects with more extreme initial scores appear to improve even when the treatment had not been successful (Westen & Morrison, 2001). (3) A group averaged raw change score that is significantly larger than that of a control group obscures whether the change is large in a few individuals or smaller in many of the individuals (Westen & Morrison, 2001). (4) A statistically larger raw change score does not necessarily imply a clinically significant difference. In a sample with little variability a reliable but small change may occur that is unimportant by anyone’s subjective judgment (e.g., in a treatment for obesity, treated subjects on average lose five pounds more than untreated subjects). (5) Raw change scores say nothing about the average post-treatment level of pathology and whether this lies in the normal range. (6) Raw change scores offer no standard technique for combining separate measures in a composite change score, as sensitivity to change varies greatly among measures (Lambert & Hill, 1994).

The last two original methods used to assess change were statistical modifications of the raw change score method, introduced to address a few of the shortcomings listed above. Investigators switched from an ANOVA of raw change scores to an analysis of covariance (ANCOVA) of post-treatment scores, with the pre-treatment score as a covariate (Lambert & Hill, 1994). This was an attempt to minimise the bias of pretreatment score and regression to the mean, but did not address any of the other

questions. More sophisticated researchers tackled the same problems using linear regression to predict a post-treatment score on the basis of regression to the mean and then analyzing differences among how observed scores differed from expected values (called “residualized change scores”) among study groups (Lambert & Hill, 1994). Neither of these techniques was completely standardised and no studies were performed to evaluate how results using these methods compared to results from studies using raw change or percent change metrics.

Based on the shortcoming of previous techniques, it is possible to construct a list of desired attributes for a method that captures clinically significant change. Such a method must start by using a standardised technique to demonstrate statistical reliability of change and, if that condition is met, be able to quantify whether that change is clinically significant based on a set of established norms. A method must exist for using these norms to describe different levels of improvement, ranging from mild improvement to complete recovery. The technique should be applicable to either group means or to individuals, depending on the level of analysis desired. Ideally the methods would provide a way of avoiding the bias of pre-treatment differences and regression to the mean. The larger goals of a method for assessing clinically significant change cannot be met with only one measure, but require data from multiple perspectives on different constructs including symptom change, meeting role demands, functioning in every day life, quality of life, and subjective judgments of client, therapist, and significant others (Kazdin, 1994; 1999). Therefore, a good technique must have a standard system for treating such variables similarly and, if desirable, combining them into a composite measure.

While pursuing a method that satisfy these criteria, it is also important to recognize problems that are inherent to the field or have not yet been addressed. First, we must accept that a method for assessing clinically significant change can only be as good as the

outcome measures on which it is based (Jacobson et al., 1999; Kazdin, 1999). In particular, our methods will depend heavily on the psychometric properties of these measures and norms for normal and dysfunctional populations. Therefore, we depend on measures that are widely enough used to have well established psychometric properties and norms, such as those described earlier in this chapter. All measures have biases, and existing symptom measures are often criticised for their heavy weighting toward pathology, leading to floor or ceiling effects when applied to normal or improved populations (Lambert & Hill, 1994). Available norms for healthy populations have most often been collected from a single assessment, and it is problematic that these are then being compared to scores from study samples who are assessed multiple times (Kazdin, 1999). Scores in the normative range may have different meaning in a community sample than they do in a sample that was previously dysfunctional. For both these reasons, non-treatment control groups are important for the assessment of clinical significance. Finally, due to a justifiable focus of the field on first establishing methods for assessing significant improvement with single measures, not enough work has been done yet to assess deterioration or to deal with multiple measures (Jacobson et al., 1999). Hopefully, advances in these areas will follow.

Current state of the art methodology for capturing clinically significant change in psychotherapy research is divided into four components: (1) use of a reliable change index, (2) comparison with normative and dysfunctional groups, (3) subjective and quality of life evaluations, and (4) assessment of social impact. Advances in the first two categories were pioneered by Jacobson in his classic 1984 (Jacobson, Follette, & Revenstorf, 1984) and 1991 (Jacobson & Truax, 1991) papers, with revisions and discussion by other researchers including Christensen and Mendoza, Kendall, Speer, and Hsu. Subjective, quality of life, and social impact evaluations each have their own literatures and are relatively recent entries into the field of clinically significant change.

All four components are well described in recent reviews in the *Journal of Consulting and Clinical Psychology* (June 1999) and *Behaviour Research and Therapy* (December 1999).

Reliable change index (RCI)

In 1984, Jacobson and colleagues (Jacobson et al., 1984) suggested a method for converting the raw difference between pre-treatment and post-treatment scores for an individual on any measure into a standardised reliable change index, based on the variability of the pre-treatment and normal population scores and the test-retest reliability of that measure. If the reliable change index exceeds 1.96, this individual is said to have shown statistically reliable change on that measure, or in other words, the likelihood that this change is due to chance is less than 0.05. Christensen and Mendoza (1986) pointed out mathematical problems with the way this index was calculated and in 1991 Jacobson and Truax issued a revised formula which has since become the standard for assessment of statistically reliable change.

$$RC = \frac{x_2 - x_1}{s_1 \sqrt{2(1 - r_{xx})}}$$

where RC = reliable change index, X_2 = post-treatment score, X_1 = pre-treatment score, S_1 = standard deviation of the combined control group, normal population, and pretreatment experimental group, and r_{xx} = test-retest reliability of the measure. In order to make it easier to apply the RCI criterion to individuals, many investigators simplify the procedure by calculating a single score (which is, confusingly, also referred to as a reliable change index, but we will call RC') against which a raw change score can be compared directly.

$$RC' = 1.96 \times s_1 \sqrt{2(1 - r_{xx})}$$

If a raw change score is greater than RC' the criterion for statistically reliable change is met. Although, as initially presented, reliable change was intended to be a method for

documenting reliable *improvement* there is no statistical reason why it cannot be used to demonstrate reliable deterioration as well, and has been used by a few investigators for this purpose (Lunnen & Ogles).

Several attempts have been made to further improve the RCI criterion. Hsu (1995; 1999) suggested modifying the formula for the RCI in order to more conservatively take into account the possibility of regression to the mean (this is sometimes referred to as the GLN method, in comparison with the JT method of Jacobson & Truax). Speer (1992; 1999; Speer & Greenbaum, 1995) proposed the alternative Edwards-Nunally method which allows greater sensitivity to change in measures with poor reliability. Finally, Hageman and Arrindell (1999a; 1999b) suggested a way to refine the JT method by calculating tailored cutoffs for each individual (known as the HA method). There has been some attempt to compare these techniques (see Behaviour Research and Therapy, December 1999), and some researchers have concluded that threats from regression to the mean are exaggerated (Speer, 1999). At this point, the Jacobson 1991 formula remains the standard for determination of statistically reliable change.

The shortcomings of the RCI are straightforward. Alone, it assesses only the likelihood that a change is real and says nothing about whether the change is clinically meaningful. Jacobson (1999) chides researchers who mistakenly use the RCI as a measure of clinically significant change. Second, the RCI makes no allowance for regression to the mean or the bias of pre-treatment scores. Finally, the index is designed to be a conservative assessment of reliability and may render measures with poor test-retest reliability or large variability in the pre-treatment sample useless as a measure of change, because of the large differences in score it would require to produce a reliable result.

Dysfunctional group and normative comparison

The basis for most assertions of clinically significant change lie in the comparison of an experimental population with normative data from normal and/or dysfunctional

groups. In one definition of clinically significant change, advocated by Kendall, the criteria for change is only met when treated individuals are “indistinguishable from normal individuals with respect to their primary complaints following treatment” (1999, p. 285). Since the major diagnostic schemes do not offer standard methods for making this judgment and traditional statistical tests are limited in demonstrating equivalency because they are not designed to confirm the null hypothesis, Kendall offers a new method. He suggests conducting two one-sided t-tests to show that the post-treatment group mean is neither significantly lower than an assigned cutoff value below the normative mean (δ_1) nor significantly higher than an assigned cutoff value above the normative mean (δ_2). Combined with a traditional t-test this method leads to one of four results when comparing an experimental and control group: (1) significant equivalency test and nonsignificant traditional t-test indicate that the groups are equivalent, (2) significant equivalency test and significant traditional t-test indicate that the groups are clinically equivalent but statistically different, (3) non-significant equivalency test and significant traditional t-test indicate that the groups are not clinically equivalent, and (4) non-significant equivalency test and nonsignificant traditional t-test indicate equivocal results, requiring more statistical power (Kendall et al., 1999). The most typical application of this procedure (called normative comparison) is to separately compare the pre-treatment and post-treatment scores of an experimental group with those of a normal control. If the pre-treatment and normal groups are non-equivalent, but post-treatment and normal groups are equivalent, one might conclude that clinically significant change has occurred (Kendall & Grove, 1988; 1999; Nietzel & Trull, 1988). A comparison might also be conducted between the experimental group and a known dysfunctional sample in order to demonstrate that these groups are equivalent prior to treatment and non-equivalent following treatment.

Jacobson (1984; 1999; 1991) advocates using cutoff points as a means for determining when a subject has significantly improved. If the change in a subject's score exceeds RC' and crosses the designated cutoff, one can conclude that clinically significant change has occurred. He suggests three methods for computing a cutoff point, chosen on the basis of how conservative a criterion is desired, what normative information is available, and what the distributions of normal and dysfunctional populations on this measure are thought to look like: (a) a cutoff two standard deviations away from the mean of the dysfunctional population (most stringent, criterion but requires data on the dysfunctional population), (b) a cutoff two standard deviations away from the mean of the normal population (most lenient criterion, but requires data on a normal population), (c) a cutoff between normal and dysfunctional population means mathematically designed as the point at which it becomes more likely that the individual's score comes from one distribution than the other (a good compromise but requires data on both normal and dysfunctional samples). As long as the distribution of normal and dysfunctional population scores is thought to be bimodal and overlapping, the third cutoff is the most reasonable choice (Jacobson et al., 1999). It is calculated according to the formula:

$$c = \frac{s_0 \bar{X}_1 + s_1 \bar{X}_0}{s_0 + s_1}$$

where c = clinical cutoff between normal and dysfunctional groups, \bar{X}_0 = mean of normal population, s_0 = standard deviation of normal population, \bar{X}_1 = mean of dysfunctional population, and s_1 = standard deviation of dysfunctional population.

Although Jacobson's cutoff is in widespread use in the psychotherapy literature, several investigators have critiqued it on the grounds that a single cutoff is arbitrary and is based on rarely confirmed assumptions that the distribution of measure scores is

binormal (i.e., bimodal and normal with respect to normal and dysfunctional populations) (Hollon & Flick, 1988; Tingey, Lambert, Burlingame, & Hansen, 1999; Wampold & Jenson, 1986). One suggestion has been to view the distribution on a given measure as representing several “adjacent samples” on the continuum from normal to dysfunction (e.g., very well functioning normals, normals, outpatients, and inpatients or, with respect to a measure such as the SCL-90, asymptomatic, mildly symptomatic, moderately symptomatic, and severely symptomatic) (Tingey et al., 1999). Reliable change can be said to occur if patients move from their starting group to an adjacent group. In order to be useful, such a computation would require exceptionally good psychometric properties for the measure to which it is applied. Alternatively, Jacobson (1999) has suggested that by using the RCI confidence intervals could be established around the chosen cutoff, and subjects who fall within these boundaries be labeled “indeterminate.” If too many subjects are labeled indeterminate, though, this method could be a serious detriment to experimental power.

Several investigators have suggested that the shortcoming of using group comparisons to evaluate clinical change with any one measure, can always be addressed by applying this criterion to several measures and then study the pattern of results (Kazdin, 1999; Kendall et al., 1999). One might even consider testing group differences with multiple cutoffs or criteria, taking advantage of the advantages and disadvantages of each to construct a better picture of the results. The drawback to such an approach is that as more results accumulate, they become harder to interpret coherently, and an enterprising investigator can find evidence for any hypothesis by trying enough methods for evaluating clinical change.

Most of the problems with using group comparison for evaluating change have been described above. Existing methodology does not make it easy to decide which system to use. Cutoffs may seem arbitrary, are based on assumptions about the

distribution, and, in the case of Jacobson's 1984 cutoff c , require extensive normative data about normal and dysfunctional populations. Kendall's method of group equivalency does not specify how the δ values should be selected, can be used with either normal or dysfunctional groups, and does not allow for designation of change in individuals. Other shortcomings of these measures have been noted. This system does not take into consideration that measures designed for use in a clinical or non-clinical sample may be less reliable when applied to the other sample type. In an equivalency test, if the ratio of the variances of the two populations being compared is greater than 3:1, statistical tests are likely to be biased (if variance is proportional to sample size, tests are conservatively biased, if variance is inversely proportional to sample size, tests are liberally biased) (Kendall et al., 1999). Finally, the combination of Jacobson's RCI and cutoff criteria may be too stringent to detect clinical change. In particular, mildly symptomatic clients are unable to reach criteria due to the floor effect and very symptomatic clients may improve significantly but still not meet criteria due to the normality requirement (Lunnen & Ogles, 1998).

To date, only a few studies have been conducted to explore the relationship between different measures of clinically significant change. Ankuta and Abeles (1993) used Jacobson's 1991 criteria for distinguishing improvers and non-improvers, and confirmed that improvers reported higher perceived satisfaction than did non-improvers. However, they failed to collect data from therapist and significant other, used only a single measure of satisfaction, and did not attempt to identify clients who deteriorated during the course of therapy. Speer and Greenbaum (1995) compared the Jacobson criteria with four similar approaches for measuring clinically significant change, and found a great deal of overlap. Their findings are limited, though, by the fact that all five approaches made similar assumptions about distribution of scores. Finally, Lunnen and Ogles (1998) categorized 52 outpatients into improvers, non-changers, and deteriorators

and compared them on a range of client, therapist, and significant other measures of perceived change, satisfaction, and therapeutic alliance. They found that improvers were distinguishable from non-improvers on client and therapist measures of perceived change and helping alliance, but less so on satisfaction. Significant other reports did not differentiate between improvers and non-improvers and no measures were able to distinguish non-changers from deteriorators. These findings confirm previous observations that satisfaction and outcome are not always related and that different rater perspectives give different information (Conway & Ross, 1984; Pekarik & Wolff, 1996). Without a doubt further empirical research is needed into all the methods for assessing clinically significant change in areas of statistical reliability, group comparison, subjective assessment, and social validity. Studies need to look not only at the properties of individual measures and methods, but of techniques for combining measures across these categories and putting together a standardised science of clinical assessment.

Practical examples of composite measures of clinically significant change

In any major study of psychotherapy outcome, it is not enough to assemble an assortment of useful outcome measures and techniques for assessing change. One must go one step further in choosing a method for how to combine these measures if straightforward and parsimonious hypotheses about improvement are to be tested using statistical tests. Of all the methods discussed, only one puts forth a standard technique for calculating composite scores. Jacobson (1999) suggests that using the RCI a standardised “true score” can be calculated for improvement on a given measure that is independent of the sensitivity and psychometric properties of that measure. Multiple standardised true scores can then be averaged for a given individual, and the resulting score compared with a cutoff to determine clinical significance. Unfortunately, since the resulting measure has different properties than its component parts, cutoff points should

be established separately using normative data on the composite measure in normal and dysfunctional samples.

Given the ideal, set forth by Jacobson, it is useful to also have a sense of the practical techniques that have been employed by psychotherapy researchers for building composite measures and assessing clinically significant change. In the Penn Psychotherapy Research project (Horowitz, Rosenberg, Baer, Ureño, & Villaseñor, 1988; Horowitz, Rosenberg, & Bartholomew, 1993; Luborsky et al., 1980) raw outcome data consisted of five patient self-report measures, four therapist report measures, and four independent evaluator measures, all collected at multiple time points in order to assess change. Within each evaluator (patient, therapist, and independent evaluator) raw change scores between two time points were converted to residual change scores, then to z-scores, and averaged, producing three composite measures of improvement. These measures were then converted to z-scores again and averaged to form a single composite measure. Though the authors justify the use of residual change scores as a way to minimise the bias of pre-treatment scores, and cite methodological references which do the same (Fiske et al., 1970; Manning & DuBois, 1962; Mintz et al., 1979) no practical procedure for calculating residual change scores is ever outlined.

Other techniques for measuring change have been less opaque but also less sophisticated. Mintz (1981) represents a great deal of psychotherapy research in using subjective ratings by clinical and non-professional raters (ratings of the two correlated well) to classify patients according to their clinical material as not improved, slightly improved, moderately improved, or much improved/recovered. A study of inpatient psychodynamic therapy in Stuttgart (Fonagy et al., 2001) evaluated four improvement criteria for each patient (meeting of therapeutic goals, change on a self-report symptom questionnaire, improvement of self-reported general well being, and improvement of self-reported capacity to work) and then reported the percentage of patients in the

sample who met 0, 1, 2, 3, or 4 of these criteria. This random sampling of techniques represents where much of psychotherapeutic research is, and reminds us of the need for rigorous standardised techniques.

Summary

The introduction in this section to issues of clinically significant change has important implications for how change should be conceptualized, measured, and reported in good psychotherapy research. First, it is clear that a standardised technique must be applied for demonstrating that change is statistically reliable and clinically meaningful. Although not perfect, Jacobson and Truax's 1991 recommendations are the best existing methods for these purposes and should be used with attention to the recommendations and caveats in the literature (e.g., application of both RCI and cutoffs, examination of measure distributions, and careful measurement of psychometric properties). It is also clear that since no single measure of change now satisfied all our requirements, it is reasonable to use several such measures, with multiple measures, perspectives, and even cutoff standards to get a sense of whether these methods agree. Third, researchers need to make more use of subjective and social impact measures, when possible, to supplement, but not replace, the information gained from more standard measures. Fourth, though almost no standard techniques exist for creating composite measures, it is worthwhile in many studies to follow Jacobson's recommendations or explore new means of calculating such a score. Finally, it is clear that more empirical research is needed to validate methods for clinically significant change. Clinical change metrics are no better than the measures on which they are based, therefore, increased standardisation with more psychometric data is needed on individual measures.

2.5 Conclusion

The preceding chapter reviews a broad range of issues and techniques within current psychotherapy research methodology, with an eye to informing the construction of a good study. It is tempting to conclude that with the enormous number of study designs, measures, and measures of change and the seemingly infinite number of combination of these, the psychotherapy researcher has no chance of thoughtfully and scientifically building a study. While it is true that the field is sorely lacking in standardised methods for choosing methodology (as well as for the methodologies themselves), we believe it is possible, through familiarity with the literature described above, to make good choices. For study design, researchers should base their selection on the questions they wish to ask, their available resources, and the extent to which they are able to intervene in the treatment of the population in question. RCT's still hold the promise of the best design, but can be reasonably replaced by quasi-experimental studies when these are more practical, as long as alternative explanations for findings are carefully considered. Process measures and standardised indices of reliable change should accompany every study and will add greatly to the interpretability of the results. Finally, with the large number of existing outcome measures, the aim of the experimenter should be to balance his selections between the need for measures that answer the questions he poses and the important of using a standard battery that is shared by other psychotherapy studies.

CHAPTER 3. DESCRIPTION AND ASSESSMENT OF YOUNG ADULT SAMPLE AND TREATMENT USING STANDARDISED MEASURES

3.1 Introduction

The usefulness of a psychotherapy research project depends, as in any scientific field, on the degree to which the investigators approach interesting questions with the appropriate design and methodology refined by previous studies. Based on the psychoanalytic research presented in Chapter 1, we begin broadly and ambitiously with six fundamental questions: (1) Is psychoanalysis effective?, (2) How do psychoanalysis and psychodynamic psychotherapy differ?, (3) What patients are more likely to benefit from psychoanalysis?, (4) What makes psychoanalysis effective?, (5) What is the process of psychoanalysis?, and (6) What do we know about psychoanalytic treatment of young adults? Psychoanalytic research over the past 87 years has shed light on all of these questions, particularly the first three, but invites the replication of findings, resolution of contradictory results, and increasing refinement of the answers.

3.1.1 *Study design*

The range of study designs and modes of measurement, as presented in Chapter 2, is broad and offers the possibility of customizing the methodology of a study for the particular questions on which the investigator chooses to focus. The study described in this thesis began with an interest in the process and efficacy of psychoanalysis and psychodynamic psychotherapy in young adults which was then matched to a set of design constraints. First, we began with the belief that these questions require an abundance of frequently collected process and outcome data in long-term treatments and concluded that for this to be manageable the sample size would have to be relatively small. Second, since the study started with a group of patients already in psychoanalysis and was only subsequently broadened to include patients in psychodynamic psychotherapy, randomizations of subjects into treatment was not possible, and we

settled for the next best alternative, assigning the first set of subjects to psychoanalysis and the second set to psychodynamic psychotherapy, while collecting enough background data to check whether the groups were comparable. Third, based on a theoretical emphasis on structural change and the intricacies of process (as described in Chapter 2), it was decided to go beyond an established core-battery of measures and introduce new methods in these areas.

The need for study designs that respect these constraints is the subject of a growing literature on “quasi-experimentation” (Cook & Campbell, 1979; Kazdin, 1994). Quasi-experiments are studies in which, due to practical considerations, the conditions of true experiments or RCTs are approximated but not fulfilled completely (Kazdin, 1994). Cook and Campbell define quasi-experiments more specifically as “experiments that have treatments, outcome measures, and experimental units, but do not use random assignment to create the comparisons from which treatment-caused change is inferred. Instead, the comparisons depend on nonequivalent groups that differ from each other in many ways other than the presence of a treatment whose effects are being tested” (1979, p. 6). They divide quasi-experimental design into four categories: (1) designs that often do not permit reasonable causal inferences (one group with pretest and posttest measures, or one or more nonequivalent groups with posttest only measures), (2) nonequivalent control group designs that are generally interpretable (untreated control group design with pretest and posttest measures, removed, reversed, or repeated treatment design, and cohort designs in institutions with cyclical turnover), (3) interrupted time series design (with or without a control group and removed or replicated treatment), and (4) correlational design. In the Open Door Review (ODR) of Outcome Studies in Psychoanalysis Fonagy (2001b; 2001) refers alternately to quasi-experiments and “open trials” which he uses to characterize studies with methodology in between that of RCTs and single-case or observational studies. The best known of the quasi-

experimental studies described in the ODR is the Stockholm Outcome of Psychotherapy and Psychoanalysis (STOPP) project (Grant & Sandell, 2004).

3.1.2 Measure selection

To have a meaningful dialogue with existing psychoanalytic and psychotherapy literature it is necessary to begin with a standardised battery of outcome measures. In fact, one of the principal criticisms of past research has been that measures are created and discarded with such abandon that it is unnecessarily difficult to compare findings from different studies (Kazdin, 1994). As discussed in Chapter 2, a core battery of standardised symptomatic and diagnostic measures were selected for the evaluation of response to treatment in this study. Also as argued in Chapter 2, a statistically reliable and clinically meaningful measurement of change depends not only on the measures chosen, but on the statistical method used for making this determination. A technique for taking this into consideration was developed in this study and is described in the Methods section below. Finally, to go beyond a replication of past studies and to introduce new methods for understanding process and structural change, two new measures are introduced, one (the Young Adult Weekly Rating Scale) created *de novo* for this study, and another (the Adult Attachment Interview and coding scheme) imported from the developmental psychology literature. Their application is described in later chapters.

3.1.3 Literature review on basic questions addressed in this chapter

The information collected by any study of psychotherapy process or outcome can be divided into three broad categories: (1) initial description of the subjects before they undergo an intervention, including background demographics, psychological capacities, psychological distress, and diagnoses; (2) description of the treatment which they undergo; and (3) measurement of change, either during therapy or after therapy has ended. A great deal of complexity and several decades of literature are associated with each of these categories. Research on psychological tests and measurement of

psychopathology is relevant to the first category, research on the measurement of psychotherapeutic process relates to the second, and the recent and rapidly growing literature on measurement of psychotherapeutic change deals with the third. Answers to any questions about psychotherapy hinge on how we deal with the problems that each of these categories pose (Kazdin, 1994). In this chapter, we will draw together some standardised and relatively simple means for collecting information on sample, therapeutic intervention, and therapeutic outcome and use them to describe the Young Adult sample.

Demographics

Demographic variables are potentially important both in understanding the differential effects of psychotherapy among members of a treatment group, and in order to convincingly generalise the findings from a research sample to a larger population. Astonishingly, a majority of psychotherapy studies do not include information about economic status, race, and education of participants and some do not even include information as to sex, age, and type of dysfunction (Kazdin, 1994).

Many studies have found associations between demographic variables and the likelihood of a subject to continue in psychotherapy. A relatively uniform finding is that subjects with lower social class (as measured typically by the Hollingshead index in the United States or the Registrar General's Classification of Occupations in Great Britain) are more likely to terminate therapy prematurely and have a shorter average length of stay in therapy (Garfield, 1994). There is some evidence, albeit less robust, that lower educational level also predisposes to early dropout, though this varies between different studies and samples (e.g., two studies of children have shown no relation between educational status and dropout rate). Of note, the results of many studies reporting dropout rates may be poorly applicable to general practice because of the effort expended in keeping subjects (Garfield, 1994). Most discouragingly, studies have found a

tendency for clinicians to rate lower class patients as being less “psychologically minded” and to accept them less often into psychotherapy. As a general rule, patients accepted for long-term psychoanalysis or psychodynamic psychotherapy tend to be of higher social class and better educated, but it is not clear to what extent this is due to the bias of therapists or simply to the availability of such therapy for those who have the means to pay for it and are in a social milieu where it is more common (Garfield, 1994). In most studies of adults, age and sex are not significantly associated with length of stay in treatment. In adolescents, a recent study found that in a sample of subjects between 12 and 24 years old, younger age was the most significant predictor of early dropout (Baruch, Gerber, & Fearon, 1998).

Initial assessments

The consensus in the psychological test and psychotherapy research literature is that a comprehensive assessment of a subject requires an array of measures which sample multiple symptom or behavioural attributes, use multiple informants, use multiple approaches for collecting data, and are repeated as often as is practical (see Chapter 2). Despite emphasis on the different information supplied by each approach, in any clinical subject there is a great deal of overlap between measures and construct validity (i.e., the tendency for a measure to pick up only one trait) is poor (Lambert & Hill, 1994). Therefore, it is to be expected that intercorrelations between initial assessments of psychotherapy patients are high, and may make it difficult to discern patterns of symptoms, particularly in small samples, such as our own.

Some associations have been shown between initial assessment of psychotherapy patients and their tendency to drop out of treatment. In adults, subjects who remained in therapy longer were found to be more anxious, more self-dissatisfied, more willing to explore problems, more persistent and dependable, and less likely to have a history of antisocial acts (Garfield, 1994). Among adolescents, subjects who stayed in therapy had

fewer externalising problems, fewer school problems, and were less likely to be diagnosed with conduct disorder (Baruch et al., 1998).

Treatment parameters

A variety of treatment parameters are typically reported in psychotherapy studies, ranging from type of therapy and length of treatment in almost all studies, to measurements of therapeutic bond and details of the interaction in more process oriented studies. In this chapter we evaluate treatment only according to intensity of treatment (intensive vs. non-intensive), length of therapy (from first to last session), approximate percentage of available sessions attended by the patient, approximate total number of sessions attended, and number of assessment interviews (conducted by a non-treating psychiatrist) attended by the patient.

Our sample is unusual in that because of the sequential system for assigning patients to intensive and non-intensive treatment groups (see Methods section below), an association between severity of pathology and intensity of treatment is not expected. Because of the small size of the sample, there is no guarantee that significant differences between samples would not appear by chance (Kazdin, 1994) and this lack of association must be verified empirically. No association should be found between intensity and demographic variables either. Treatment attendance and length of treatment are important variables in that they are presumed to have a direct impact on therapeutic outcome, yet are often overlooked because they are difficult to study. Most research reveals that premature termination is more common than is widely appreciated and the number of actual sessions is far fewer than is deemed optimal, even in successful therapies (Garfield, 1994).

Treatment outcome

Longitudinal research on psychoanalysis and psychodynamic psychotherapy generally concludes that subjects do gradually get better over time, as indicated by

improvement on a variety of outcome measures (Sandell, 2000; see Chapter 1). Meta-analysis of a broad range of psychotherapy studies has shown that the relationship between improvement and time in therapy resemble as dose-response curve (Howard et al., 1986). The evidence that different types of therapy lead to different outcomes is scarce, but there is evidence that psychoanalysis has more immediate and more longer lasting effects than psychotherapy (see Chapter 1).

There have been no conclusive demonstrations of differential therapeutic outcome in subjects of different social class, education, or race. However, such studies are limited by the relative homogeneity of subjects in psychotherapy, particularly psychoanalysis. Research on the effect of age on psychotherapy outcome is complicated by the many different age groups into which subjects can be divided. In general, the presumption that older subjects are more rigid and fixed in their ways and therefore less amenable to psychotherapy has not been well-supported. Data is not available on the effect of age on psychotherapeutic outcome in adolescents. Sex of patient and therapist has not been shown to be a factor in psychotherapeutic outcome. The relationship between IQ and therapeutic outcome was investigated seriously several decades ago with some suggestion that subjects with higher IQ respond better to therapy, but results were rarely replicated and interest in this as a research question has since waned (Garfield, 1994). Luborsky (1980) found no relationship between Wechsler Adult Intelligence Scale and psychotherapy outcome.

A great deal of attention has been paid in the last decade to investigating the differential response of distinct disorders to psychotherapy. In a comprehensive review of psychotherapy research, Roth and Fonagy (1996; in press) divide the evidence for efficacy of different forms of psychotherapy into chapters on standard clusters of disorders. They conclude with a summary of those therapies that have been empirically validated for individual disorders. For psychodynamic psychotherapy, they list

depression, post-traumatic stress disorder, personality disorders, and generalised anxiety and psychophysiological disorders in children, as problems for which there has been some reliable demonstration of therapeutic efficacy. On a theoretical level, Strupp, Schacht, and Henry (1988) introduced the principle of Problem-Treatment-Outcome congruence which states that in psychotherapy research a patient's problems, treatment processes, and treatment outcome should all be conceptualized and measured according to a common metric. This study continues in the tradition by considering initial assessments alongside treatment parameters in attempting to predict psychotherapy outcome.

3.2 Methods

3.2.1 Sample description

Of the 26 patients entered into the research scheme, 15 began intensive therapy between August 1990 and January 1995 and 11 began non-intensive therapy between April 1994 and September 1998. All patients received some psychotherapeutic treatment. One of the 15 patients in intensive therapy was subsequently removed from the research scheme because of the patient and analyst's concerns as to the confidentiality of his material. The remaining 14 intensive and 11 non-intensive cases are listed in Tables 3.1 and 3.2 and are described in this section. Due to difficulties recruiting subjects for assessments and lapses in forms being filled out by some of the analysts, the subjects will be studied in seven overlapping groupings, each appropriate for a particular analysis. Analysis of variance will be used to confirm that demographics and initial psychiatric measures do not differ by group. For these purposes, "assessment" refers to at least some part of the psychological and psychiatric battery of measures administered to the subject by a non-treating psychiatrist, as is described section below.

Group 1 = Subjects who completed an initial assessment (n=25)

Group 2 = Subjects who completed initial and termination assessments (n=19)

Group 3 = Subjects who completed initial and termination assessments, and at least one follow-along assessment during treatment (n=10)

Group 4 = Subjects who completed initial and termination assessments, and at least one post-termination follow-up assessment (n=13)

Group 5 = Subjects who completed initial and termination assessments, and whose treatment was documented by the analyst using the YAWRS (n=18)

Group 6 = Subjects who completed initial and termination assessments, and underwent at least two coded adult attachment interviews (n=18)

Group 7 = Subjects who completed initial and termination assessments, and underwent at least two coded adult attachment interviews, including at least 1 conducted post-termination (n=7)

<i>Subj</i>	<i>Gender</i>	<i>Age at start tx</i>	<i>Tx Duration</i>	<i>Grp 1</i>	<i>Grp 2</i>	<i>Grp 3</i>	<i>Grp 4</i>	<i>Grp 5</i>	<i>Grp 6</i>	<i>Grp 7</i>
A	M	22	8 yrs	✓	✓	✓		✓	✓	
B	M	26	1.5 yrs	✓	✓			✓		
C	F	22	8 yrs	✓	✓	✓	✓	✓	✓	
D	M	21	8 yrs	✓	✓	✓	✓	✓	✓	
E	M	25	7 yrs	✓	✓	✓		✓	✓	
I	F	27	1 yr	✓						
M	F	23	2.5 yrs	✓	✓		✓	✓	✓	✓
O	M	23	7.5 yrs	✓	✓	✓		✓	✓	
P	M	25	4 yrs	✓	✓	✓	✓	✓	✓	✓
Q	F	22	3.5 yrs	✓	✓		✓	✓	✓	✓
T	F	23	5 yrs	✓	✓	✓	✓	✓	✓	
U	M	22	5.5 yrs	✓	✓	✓	✓	✓	✓	
V	F	19	6 mos	✓						
X	F	22	5 yrs	✓	✓	✓	✓	✓	✓	

Table 3.1. Patients in intensive treatment

<i>Subj</i>	<i>Gender</i>	<i>Age at start tx</i>	<i>Tx Duration</i>	<i>Grp 1</i>	<i>Grp 2</i>	<i>Grp 3</i>	<i>Grp 4</i>	<i>Grp 5</i>	<i>Grp 6</i>	<i>Grp 7</i>
F	M	24	9.5 mos	✓	✓		✓	✓	✓	✓
G	F	23	2 mos	✓						
H	F	24	6 mos	✓	✓		✓		✓	✓
J	F	23	6 mos	✓						
K	M	18	6 yrs	✓	✓	✓		✓	✓	
L	F	25	2 yrs	✓	✓		✓	✓	✓	✓
N	F	20	3 yrs	✓	✓		✓	✓	✓	
R	F	22	1.5 yrs	✓	✓		✓	✓	✓	✓
S	F	25	6 mos	✓						
W	F	19	1 yr	✓						
Y	F	20	2.5 yrs	✓	✓			✓	✓	

Table 3.2. Patients in non-intensive treatment

Gender

Subjects were admitted to the study without any attention to gender. Of the 25 total subjects, eight were male (six intensive, two non-intensive) and 17 were female (eight in intensive and nine in non-intensive therapy). The gender breakdown in groups 2 through 7 was (male/female): group 2 = 8/11, group 3 = 6/4, group 4 = 3/10, group 5 = 8/10, group 6 = 7/11, group 7 = 2/5. Chi-square analyses showed a significant statistical association between gender and group membership for group 3 only (Chi-squared [df=1] = 6.0, Fisher's exact test $p < 0.05$).

Age

Broad age boundaries for young-adulthood were used in this study, as many clinicians have argued that the psychological challenges of adolescence and young-adulthood extend later than was previously believed (e.g., Giedd et al., 1999). Subjects ranged in age between 19 and 27 years old at the start of therapy, with a mean of 22.8 (SD=2.1). Mean ages for groups 2 through 7 were: group 2 = 22.8 (1.8), group 3 = 22.6

(1.7), group 4 = 22.9 (1.6), group 5 = 22.7 (1.8), group 6 = 22.7 (1.8), group 7 = 23.7

(1.5). Analysis of variance (ANOVA) revealed no statistical association between age and group membership.

Socioeconomic status (SES)

SES was assigned to each subject using the Registrar General's Classification of Occupations according to the highest level at which the subject, subject's father, or subject's mother had worked. Eight subjects were in class I (highly skilled professionals), 11 in class II (professionals with lower degrees), none in class III (shopkeepers and skilled craftsmen), none in class IV (semi-skilled, clerical, or secretarial labor), and two in class V (unskilled labor). SES information was unavailable for four subjects. There was no statistical association between SES and group membership. Of the subjects themselves, 11 were students, nine were unemployed, and five had jobs. The mean SES for intensive subjects was 1.5 (SD=.52, n=13) and for non-intensive subjects 2.6 (SD=1.5, n=8). No association was found between SES and group membership.

Intelligence (IQ)

Subjects were assessed at intake using the National Adult Reading Test (NART), an oral test of word pronunciation designed to estimate premorbid IQ in demented patients but also useful as a quick and rough estimate of IQ in normal and psychiatric populations (Nelson, 1982). Subjects were given a list of 50 words, in increasing order of pronunciation difficulty, and asked to pronounce each. The examiner recorded how many errors were made by the subject and this number was used to calculate estimates of full-scale, verbal, and performance IQ using a simple linear regression formula.

Preliminary studies have found that the ability to pronounce words is highly correlated with WAIS Full-Scale IQ ($r=0.75$) (Nelson, 1982). Intensive subjects ranged in IQ from 117 to 125 (mean=120.6, SD=2.9, n=13). Only one non-intensive subject received the

NART and received a score of 125.2. No association was found between IQ and group membership.

Psychiatric history

Psychiatric history was collected through interviews by clinicians before the subjects were accepted into the research scheme and separately from psychiatric diagnoses assigned in later assessment interviews. Of the 23 subjects for whom psychiatric history was available (records of interviews with G and S were unavailable), none had previous psychiatric hospitalizations. Four subjects (H, L, N, and R) had been on psychotropic medications but none were during the course of the study. All but three subjects (M, U, and W) had received previous outpatient psychotherapy for various lengths of time. Four had made at least one suicide attempt (A, H, G, and V) and an additional seven had histories of recurrent suicidal thoughts and/or had made at least one serious suicidal gesture (D, E, J, M, N, T, and Y). Four subjects presented with a history of violent behaviour (C, E, O, and U). Three subjects had long-standing medical problems: two with poorly controlled insulin-dependent diabetes mellitus (D and I), and one with inflammatory bowel disease (J). Finally, although it was not assessed specifically, four subjects presented with histories of sexual abuse (E, N, V, and W). No element of psychiatric history was found to be statistically associated with membership in any of the seven subject groupings, though the small number of subjects limited the usefulness of these analyses.

Associations between demographic variables

No significant associations were found between gender, age, SES, and IQ in this sample.

3.2.2 Subject recruitment and assessments

All subjects were presented to the Young Adult Research Group by a personal referral to one of the members of the group or by a general referral from the Anna Freud

Centre. Subjects then underwent an unstructured psychoanalytic assessment interview, lasting approximately 75 minutes, with a member of the group (in 18 out of the 23 interviews with available information, the interview was performed by DM, a male psychiatrist member of the group). In most cases, the interviewer saw the subject for one or more follow-up interviews, until he/she was able to feel confident about judging the subject's suitability for inclusion in the study (number of follow-up interviews for the 23 subjects with available information ranged from 0 to 6, with a mean of 1.4, and a median of 1). In all but one of these cases, the initial interview was followed by a psychoanalytic assessment interview with a member of the group of the opposite sex as the first interviewer. In 12 of the 23 cases this second interviewer was the same person who later became the treater. Following each interview, the analyst wrote a three to five page narrative account, focusing on the patient's current situation, family situation, perceived desire for treatment, willingness to commit to the requirements of the treatment and research protocol, and, the analyst's judgment of suitability for analytic treatment (based roughly on psychological mindedness and patient's presenting symptoms). After all interviews were completed, the subjects were then presented at a meeting of the Young Adult Group and a decision as to whether to accept the patient into the scheme was arrived at through discussion. Using this system fewer than five subjects were deemed unsuitable for the research scheme, several because they were unable to make the time commitment required, and one because it was deemed clinically advisable for him to have a male therapist, but none had time available. No other information is available about the unsuitable subjects. Once a subject was approved, he/she was offered as a patient to anyone in the group with an available opening in their schedule. No attempt was made to match specific patients with therapists because of clinical interest, gender, compatibility, or any other reason.

Initial psychiatric/psychological assessment

After being accepted into the study, subjects were scheduled to meet for two two-hour interviews with one of two psychiatrist members of the young adult group (AH for patients accepted for intensive treatment, and AB for patients accepted for non-intensive treatment). This psychiatrist had not been present at the meeting when the subject's entry into the scheme was discussed and was blind to results of previous evaluations. In the first meeting, the psychiatrist administered and audio-taped a semi-structured interview about the history of the subject's attachment relationships (AAI). In the second meeting, subjects were asked to fill out six self-report questionnaires assessing symptomatology, functioning, and personality (SCL-90-R, BDI, STAI, SAS-M, EPQ, and SCID-IIQ), were administered an oral IQ test (NART), and were questioned in two semi-structured psychiatric diagnostic interviews during which the psychiatrist filled out a paper-based rating questionnaire (SADS-L and SCID-II). Following the psychiatric interviews, the assessing psychiatrist filled out one general paper-based psychiatric diagnostic questionnaire (BPRS). Finally, the psychiatrist questioned the patient in a diagnostic interview specific for borderline personality disorder and filled out a corresponding paper form (DIB).

Follow-along and follow-up assessments

Eighteen months following the initial assessment, and every 18 months thereafter for as long as treatment continued, subjects were written to or telephoned with scheduled times for follow-up interviews. These were conducted in one two-hour session, during which the subjects were audio-taped completing the AAI and answering open-ended questions regarding their current life, and during the previous and next 18 months, filled out five self-report questionnaires (SCL-90-R, BDI, STAI, SAS-M, and SCID-IIQ), and were questioned in two psychiatric diagnostic interviews (SADS-L and SCID-II). EPQ, BPRS, NART, and DIB measures were not repeated at follow-up

assessments because they measure mostly stable dimensions, not believed to be responsive to psychotherapy, and because one session instead of two was necessary to make scheduling possible. As soon as psychotherapy was terminated, either with or without the agreement of the therapist, a follow-up assessment was scheduled, followed-up by further assessments at 18-month intervals. These interviews have been continued for up to five years after termination of the treatment.

Because of difficulties contacting patients and arranging a time in which they were available to meet with the psychiatrist assessor, the intervals between assessments, and even the occurrence of assessments at all, differed widely from what was planned. The initial assessment was held a mean of 0.6 months before the patient's first therapy session (range=6.7 months before the first session to 5.9 months after, SD=2.9, n=24), not including an outlier of one patient (subject S) who, because of her own scheduling problem, did not begin therapy until two years after the initial assessment. Twenty out of 25 initial assessments took place before the first therapy session. The 19 post-termination interviews conducted took place a mean of 1 month (SD = 4.9) after the final therapy session, though due to difficulty in scheduling these sessions and some last minute uncertainty in when the treatment was terminating, the timing of this assessment ranged from 8.7 months before to 9.5 months after the last psychotherapy session. Fifteen of the 19 interviews took place in a time period ranging from 2 months before to 5 months after the last psychotherapy session. The 25 follow-along interviews (conducted during treatment) took place a mean of 22.4 months after the previous assessment (range 11.7 to 36.6 months, SD=5.7) and the 18 follow-up interviews took place a mean of 25.9 months after the previous assessment (range=9.6 to 40 months, SD=7.8) not including one outlier of a subject who could not be re-assessed until 66 months after the previous interview because of geographic relocation. The large range in timing of follow-along and

follow-up assessments was due to subjects who were either difficult to locate or repeatedly missed or canceled meetings with the assessing psychiatrist.

3.2.3 Measures

SCL-90-R

The Derogatis SCL-90-R (Symptom Check-List) is a 90-item self-report inventory which asks subjects to assess psychological symptoms of distress on a 5-point scale of discomfort based on their experience of each symptom during the previous week. Nine symptom dimensions (somatization, obsessive-compulsive, depression, anxiety, hostility, phobic anxiety, paranoid ideation, and psychoticism) and three global indices (global severity index, positive symptom total, and positive symptom distress index) are calculated from the responses. The questionnaire is introduced briefly to the subject, who usually takes 12 to 15 minutes to complete the entire form.

The SCL-90-R was published in 1977 and has since been used widely throughout the psychiatric and psychological literature. SCL-90-R has been used in psychiatric and medical populations to quantify symptomatology in virtually every type of psychological disorder for the purposes of both clinical evaluation and research. Norms have been published for many of these populations (Derogatis, 1977; Derogatis & Cleary, 1977). The only subjects considered poor candidates for administration of the SCL-90-R are those who are delirious, retarded, floridly psychotic, or motivated to distort their answers.

Studies of the psychometric properties of the SCL-90-R have shown the individual subscales and PSI to be highly reliable and valid indicators of psychopathology. Derogatis (1977) reports two studies finding high internal consistency of scales, with alpha coefficients ranging from 0.77 to 0.90 and 0.95 for the overall measure. Test-retest reliability ranged from 0.78 to 0.90 for a 1-week interval and 0.68 to 0.83 for a 10-week span. Validity studies have shown that factor analysis supports the subscale structure of

the test, factor structure is invariant to different subject populations, and that SCL-90-R scales correlate well with other validated measures of psychopathology such as the MMPI, the Hamilton Rating Scale for Depression, the General Health Questionnaire, and other common measures. Studies have also found that the SCL-90-R sensitively and specifically distinguishes cases from controls for anxiety and depressive disorders, post-traumatic stress disorders, suicidal behaviour, drug and alcohol abuse, physical and sexual abuse, and sexual dysfunction (Derogatis, 1977).

BDI

The Beck Depression Inventory (BDI) is a 21-item self-report inventory which assesses the severity of a subject's complaints, symptoms, and concerns related to his/her current level of depression. It is one of the most widely used instruments for the assessment of depression in psychiatric patients, as well as in normal populations, and is estimated to have been used in over 1,500 studies by 1985.

The items of the BDI are rated from 0-3 in terms of severity; they evaluate mood, pessimism, sense of failure, self-dissatisfaction, guilt, punishment, self-dislike, self-accusations, suicidal ideas, crying, irritability, social withdrawal, indecisiveness, body image change, work difficulty, insomnia, fatigability, loss of appetite, weight loss, somatic preoccupation, and loss of libido. The measure takes approximately 10-15 minutes to complete, and is scored by the researcher by summing the 21 items, with a maximum total of 63 (Beck & Steer, 1993).

Beck (1993) reports good internal consistency of the BDI items from a meta-analysis, with a mean alpha coefficient of 0.86 for psychiatric and 0.81 for non-psychiatric samples. Test-retest reliability is highly dependent on the interval between measurements, particularly if treatment is administered between measurement points. Beck (1993) cites a meta-analysis which found reliability varying between 0.48 to 0.86 in psychiatric samples, and 0.60 to 0.90 in non-psychiatric samples. Validity of the BDI has

been confirmed by numerous studies finding high correlations between BDI score and psychiatric patient status, clinical assessments of depression, the MMPI depression scale, the Zung Self-rating Depression Scale, and the Hamilton Psychiatric Rating Scale for Depression. The BDI has not been found to be significantly related to gender or age.

In this study, the BDI will be used to assess change in response to psychotherapy. Clinical change will be evaluated by dividing scores into non-case ($BDI \leq 16$) and case ($BDI \geq 17$) status, and then observing change between these groups. The reliable change index was calculated from a sample of 248 outpatients with mixed depressive diagnoses (Beck & Steer, 1993) having a mean BDI of 23.2, and S.D.=9.55.

STAI

The State-Trait Anxiety Inventory (STAI) Form Y consists of two 20-item self-report questionnaires, the first evaluating the transitory emotional condition of tension, apprehension, nervousness, worry, and autonomic arousal (state or “S” anxiety) and the second evaluating stable individual differences in anxiety proneness and response to stressful situations (trait or “T” anxiety). Items on S-Anxiety ask the subject to rate statements about feeling calm or anxious *today* as not at all, somewhat, moderately so, or very much so (scored 1 through 4). Items on T-Anxiety ask the subject to rate statements about how anxious or calm they *generally* feel as almost never, sometimes, often, and almost always (scored 1 through 4). As of 1983, the STAI had been used in over 2,000 studies, ranging from psychology and education to medicine and dentistry. The S-Anxiety scale is typically used to assess transitory responses to stressful experiences, while the T-Anxiety scale is used to identify persons with high levels of “neurotic” anxiety, and to keep track of changes in response to treatment programs.

The STAI takes approximately 10-15 minutes to complete and is scored by the researcher by separately summing the 20 items of the S and T scales, reversing (5-x) the scores of the 10 items of the S scale and the 9 items of the T scale that reflect anxiety-

absence. Spielberger (1983) does not provide recommended clinical cutoffs for the STAI, but does provide means for a non-clinical sample (subjects age 19-39, S-Anxiety: mean=36.3, S.D.=10.6, n=656; T-Anxiety: mean=35.9, S.D.=9.6, n=656) and a sample of neuropsychiatric patients with depressive and anxiety reactions (S-Anxiety: mean=50.7, S.D.=12.07, n=88; T-Anxiety: mean=49.8, S.D.=11.37, n=88). Using these values and the Jacobson-Truax (Jacobson et al., 1984) formula (see Chapter 2), clinical cutoffs of 44.0 and 43.4 were calculated for S-Anxiety and T-Anxiety, respectively.

Spielberger (1983) reports high internal consistency of the STAI scales, with alpha coefficients of 0.90 or higher in normative samples. In a sample of 104 college students, test-retest reliability over 104 days was 0.32 for S-Anxiety and 0.75 for T-Anxiety. Concurrent validity of the scales was demonstrated by high correlations between T-Anxiety and scores on Cattell's Anxiety Scale Questionnaire and Taylor's Manifest Anxiety Scale, as well as significant changes in S-Anxiety in response to stressful and relaxing situations such as surgical procedures, exams, and relaxation exercises (Spielberger, 1983).

In this study, STAI T scale will be used to assess change in response to psychotherapy. Clinical change will be measured by assigning scores to case and non-case status using the cutoffs calculated above. A reliable change index was calculated using college student test-retest data and clinical population sample standard deviation as 10.6 for T-Anxiety.

SAS-M

The Social Adjustment Scale – modified for use in British populations (SAS-M) is a 45-item self-report questionnaire designed to measure social functioning over the past two weeks in seven “role areas”: work outside the home, housework, social and leisure activities, relationships with extended family, relationship with spouse, functioning as a parent, and functioning in the family unit (Cooper, Osborn, Gath, & Feggetter, 1982).

The SAS-M takes 15-20 minutes to complete. It is scored by the researcher by calculating item means for the six role areas, calculating descriptive category means through factor groupings of the items into those reflecting performance, interpersonal behaviour, friction, and feelings, and by calculating an overall mean of all 45 items. A higher score on any of these scales indicates greater social maladjustment and poorer functioning. The overall adjustment mean is most commonly used to assess social functioning, particularly in response to treatment - (Zuckerman, Prusoff, Weissman, & Padian, 1980).

Weissman (1978) reports high internal consistency of the SAS self-report (alpha coefficient = 0.74) and test-retest reliability of 0.80. She also demonstrates validity by showing significant differences between each pair of the following four groups: community normals, acute depressives, alcoholics, and schizophrenics (schizophrenics scored highest, followed by alcoholics, acute depressives, and normals). SAS scores were correlated with overall symptom scores on clinician rated Hamilton and Raskin Depression measures and the self-rated SCL-90. Cooper (1982) found high interrater reliability on all SAS-M scales between self-report, interview, and husband report in a sample of 130 mothers.

EPQ

The Eysenck Personality Questionnaire is a 90-item inventory of yes-no questions which assesses subjects along three major dimensions of personality: psychoticism/tough-mindedness (P), extraversion, (E), neuroticism/emotionality (N), while also estimating the tendency for falsification (lying/social desirability; L). The scale aims not to discriminate between normal and clinical subjects, but rather to measure independent personality dimensions along continuous ranges from stability to illness (Eysenck & Eysenck, 1975). The EPQ takes 10-15 minutes to complete and is scored by the researcher who sums yes or no answers that correspond to the individual subscales. The E, P, and N scales emerge from a body of research using many measures,

including the MMPI and Guilford scales, that describes at least three major independent personality scales accounting for most of the variance in personality. Eysenck has demonstrated that there are strong hereditary links to personality (explaining over 50% of the phenotypic variance), as measured by the EPQ, and believes the dimensions to be closely linked to biological predispositions. Validation of the scales has been done through studies finding similar factor organization in many samples, and finding that certain populations have expected higher scores on individual scales (e.g., psychotics have high P scores).

Internal consistency of the EPQ scales has been demonstrated in normals and prisoner samples with alpha coefficients ranging from 0.68 to 0.88. E, N, and L scales are normally distributed in most populations, whereas P is skewed, reflecting either poor scale construction or the actual distribution of the trait. Test-retest reliabilities of the scales for a one-month interval range from 0.78 and 0.89. In this study the EPQ was filled out only at the initial assessment and therefore will only be used as a predictor of other outcome and process variables.

SADS-L

The Schedule for Affective Disorders and Schizophrenia (SADS) was developed as part of the NIMH Collaborative Program on the Psychobiology of Depression in order to reduce the information variance among investigators using the Research Diagnostic Criteria (Endicott & Spitzer, 1978), which were themselves later incorporated into DSM-III (Spitzer, Endicott, & Robins, 1978). The SADS is a 46-page form which prompts the clinician to investigate: (1) detailed descriptions of features of the current episode of illness at its most severe, (2) description of severity and manifestations of psychopathology during the previous week, (3) detailed description of past psychopathology and functioning relevant to evaluating diagnosis, prognosis, and overall severity of disturbance.

Tests of reliability have show that 90% of the 120 scaled items on the SADS had levels of agreement of 0.60 or higher when assessed separately by two clinicians sitting in on the same interview, while 82% of the items had agreement of 0.60 or higher when assessed by two clinicians in separate interviews, one to three days apart. Intraclass correlations for the summary scales ranged between 0.82 and 0.97 for the joint interview, and between 0.67 and 0.93 for separate interviews. Internal consistency was high for most summary scales, with alpha coefficients of 0.47 to 0.97 (Endicott & Spitzer, 1978). Validity of the SADS has been demonstrated by high correlations with other measures including self-report SCL-90 and Katz Adjustment scales (Endicott & Spitzer, 1978) and with the Diagnostic Interview Schedule (Hasin & Grant, 1987a, 1987b).

In this study the SADS-L was used as a guide for the assessing psychiatrist to diagnose DSM-III-R Axis I disorders in the subjects. Individual items were not always completed, preventing the use of summary scales, and calling into some question whether the reliability and validity of the measure is as good as reported with SADS-L in the literature. The psychiatrist did, however, take notes throughout the interview and a researcher then read these notes to confirm the presence or absence of individual diagnoses. Lifetime and current diagnoses were extracted in accordance with DSM-III-R criteria for each assessment at which the SADS-L was completed.

SCID-II

The Structured Clinical Interview for DSM-III-R Personality Disorders (SCID-II) is a clinician-administered semi-structured diagnostic interview for assessing the 11 Axis II personality disorders of the DSM-III-R plus the Appendix category self-defeating personality disorder. The SCID-II was first developed in 1984 by Dr. Jeffrey Jonas at McLean Hospital as a personality module for the Structured Clinical Interview for DSM-III. It was updated for DSM-III-R in 1986 and after reliability field trials were completed, a final version was published in 1990. The SCID-II was further revised after publication

of the DSM-IV to bring it into accord with new diagnostic criteria and make questions more reflective of the subject's inner experience; this version was published in 1997 (First, Gibbon, Spitzer, Williams, & Benjamin, 1997). Of the 67 SCID-II assessments done in this study, 53 used the 1990 version of the SCID-II, and 14 used a 1986 version.

The SCID-II (the 1990 version will be described, only slight modifications apply for the 1986 version) supplies the interviewer with over 130 suggested prompts and questions, designed to elicit focused information from the subject on the presence of symptoms and personality traits associated with personality disorders. At the end of the section, if a specified number (between 3 and 5) of statements were rated with a 3, the subject receives a diagnosis of that personality disorder.

With the 1990 version of the SCID-II, a SCID-II Personality self-report questionnaire was published which is designed as a screening tool to shorten the length of a SCID-II interview. The subject takes approximately 20 minutes to fill out a form, answering yes or no to 113 questions (119 in the 1997 version) which correspond closely to the interviewer prompts suggested in the full SCID-II. The interviewer can thus inquire only about the questions that the interviewee responded "yes" to on the questionnaire.

Williams and colleagues (1992) and First and colleagues (1995) measured test-retest reliabilities with an interval of two weeks on non-psychiatric ($\kappa=0.38$) and psychiatric samples ($\kappa=0.53$). More encouragingly, First and colleagues (1997) report a study by Malow and colleagues which found kappas of 0.87 for borderline personality disorder and 0.84 for antisocial personality disorder with interviews done 2 days apart. Weiss and colleagues (Weiss, Najavits, Muenz, & Hufford, 1995) found an overall kappa of 0.46 in 12-month test-retest reliability of personality disorders on 31 cocaine-dependent patients. Concurrent validity of the SCID-II has been demonstrated by relating SCID-II to multiple measures indicating decreased function, predisposition to

psychiatric illness, and diagnosis by means of other self-report and clinician-report measures (First et al., 1997).

In this study the SCID-II screening questionnaire and SCID-II interview were used following the SADS-L to assess personality disorders at initial assessment and all follow-ups. In follow-up interviews, although the clinicians screened for all personality disorders, particular attention was paid to previous diagnoses. The presence, absence, and number of personality disorder diagnoses will be used as one of the measures of outcome in response to psychotherapy.

BPRS

The Brief Psychiatric Rating Scale (BPRS) is a 16-item questionnaire which asks clinicians to rate the severity of the following symptoms: somatic concern, anxiety, emotional withdrawal, conceptual disorganization, guilt feelings, tension, mannerisms and posturing, grandiosity, depressive mood, hostility, suspiciousness, hallucinatory behaviour, motor retardation, uncooperativeness, unusual thought content, and blunted affect. The 16-item scores are summed to arrive at a total pathology score. The BPRS is one of the most widely used research instruments in psychiatry because of speed and ease of use, and broad coverage of major symptoms. (Gabbard et al., 1987).

Flemenbaum & Zimmermann (1973) found test-retest reliabilities of 0.57 for a taped interview and 0.76 for a narrative description. Inter-rater reliability for the narrative descriptions was 0.51. The wide use of the BPRS in conjunction with other measures of psychopathology has confirmed the validity of the BPRS in distinguishing clinical cases. In this study the BPRS was rated by clinicians using information gathered from SADS-L and SCID-II interviews at the initial assessment.

DIB

The Diagnostic Interview for Borderlines (DIB) is a semistructured interview composed of 123 operationally defined and scored variables that assess five areas of

functioning considered most characteristic of borderline patients: social adaptation, impulse/action patterns, affects, psychosis, and interpersonal relations. The original edition was published by Gunderson and colleagues in 1981 (Gunderson, Kolb, & Austin) on the basis of previous research into the homogenous characteristics of borderline patients. The total diagnostic score is the sum of the five scaled section scores; a score of seven or higher is suggested by the interview's authors (Gunderson et al., 1981) as a sensitive and specific indication of a borderline diagnosis.

Gunderson and colleagues (1981) report good inter-rater reliability of section total scores based on a sample of 16 psychiatric inpatients (mean intraclass $R=0.61$ for individual items, 0.75 for scored statements, 0.91 for section totals, and 0.77 for scaled section totals). Frances and colleagues (1984) report an average weighted kappa of 0.78 for diagnosis of borderline personality disorder in a sample of 76 outpatients. Frances and colleagues (1984) showed a significant relationship between DIB diagnostic score and DSM-III-R diagnosis of borderline personality disorder, as well as a sensitivity of 73% and a specificity of 80% when a cutoff of 7 is used for the DIB and compared against DSM-III-R. Soloff and Ulrich (1981) successfully used the DIB to differentiate between borderline patients and schizophrenic patients, and borderline patients and depressed patients, as diagnosed by APA Research Diagnostic Criteria. In this study the DIB was used at initial assessment to provide an alternative to the SCID-II in diagnosis of borderline personality disorder.

AAI

The Adult Attachment Interview (AAI) is a structured, semi-clinical interview, usually 45 to 90 minutes in length, in which a subject is asked a series of questions about recollections of family organization, relationships with caregivers, and experiences of trauma and loss. The interview is audio-taped and transcribed verbatim, and then subjected to rating by a trained and certified rater, focusing heavily on psycholinguistic

qualities of the subject's discourse (Main & Goldwyn, 1985/1994). By "surprising the unconscious" (Main & Goldwyn, 1985/1994) and analyzing the interviewer's linguistic response, the interview and coding system elicit state of mind regarding attachment processes and internal working models of relating. The theoretical model for attachment and the AAI, as well as psychometric issues, are discussed in Chapter 2.

3.2.4 Reporting of assessment results and statistical analyses

For the purpose of data analysis, assessments were organized into four categories: initial (within 6 months of onset of treatment), follow-along (more than one year after treatment began or before termination), termination (within 6 months of treatment termination), and follow-up (more than one year after termination). The number of assessments with complete data was sufficient for a meaningful analysis on at the initial (n=25) and termination (n=19) time points. Results will be presented from these assessments and they will be used, as described below to classify subjects as "improvers" or "non-improvers." Only four measures (BDI, STAI, SADS-L, and SCID-II) were consistently collected at initial assessment and termination so assessment of change was restricted to these measures. However, summary statistics presented below include means of all the continuous scales (SCL-90-R global severity, BDI total score, STAI state and trait total scores, SAS-M total score, EPQ E, P, N, and L scale scores, BPRS total score, DIB total score, and SADS-L and SCID-II number of diagnoses) and numbers of subjects in the clinical and non-clinical ranges for each scale.

Associations between initial assessment measures and demographic variables (gender, age, SES, and IQ) were investigated using Pearson correlations, ANOVAs, and chi-square analyses. A similar statistical analysis was used to check for associations between demographic variables and treatment parameters (intensity of treatment, length of treatment, time between first and last assessment, number of assessments, percentage of offered sessions attended, and estimated number of total sessions attended) and

between initial assessment measures and treatment parameters. Concurrent validity of the assessment measures was tested by using Pearson correlation coefficients, ANOVA, and chi-square analyses to find associations between measures collected simultaneously at either intake or termination.

For each measure, subjects were assigned to clinical and non-clinical groups using clinical cutoffs derived from psychometric data reported above (a score greater or equal to the clinical cutoff places the subject in the clinical range). Subjects were classified as clinical improvers if they moved from clinical to non-clinical groups and clinical deteriorators if they moved from non-clinical to clinical groups. In addition, subjects were classified as having achieved statistically reliable improvement if the improvement in score divided by the standard error of differences (calculated according to the formula outlined in Chapter 2) is greater than 1.96 and having achieved statistically reliable deterioration if the deterioration divided by the standard error of differences is greater than 1.96. Numbers of subjects in each of these groups are reported. Results of these analyses were used to determine a strategy for classifying individual subjects as overall improvers or non-improvers. Table 3.3 lists cutoff values for distinguishing clinical and non-clinical groups, standard errors of differences, and the sample information used to compute them for each of the assessment scales in which such indices are appropriate to this study.

<i>Scale</i>	<i>Clinical mean (SD)</i>	<i>Non-clinical mean (SD)</i>	<i>Clinical cutoff (*=computed)</i>	<i>Test-retest reliability</i>	<i>SE of differences</i>
SCL-90 global severity	1.26 (.68)	0.31 (.31)	0.61*	0.84	0.38
BDI total	23.2 (9.55)		17	0.72	7.1
STAI-State	50.7 (12.07)	36.3 (10.6)	43.0*	0.32	14.1
STAI-Trait	49.8 (11.4)	35.9 (9.6)	42.3*	0.75	8.1
SAS-M total	2.23 (.44)	1.70 (.30)	1.9*	0.80	0.28
SADS-L # of diagnoses	1.3 (1.2)		1	0.67	0.97

SCID-II # of diagnoses	0.83 (0.97)	1	0.53	0.94
DIB total		7		
EPQ-E		♂: 14 ♀: 13		
EPQ-P		♂: 4 ♀: 3		
EPQ-N		♂: 10 ♀: 12		

Table 3.3. Clinical cutoff, standard error of differences, and sample information.

The extent to which change scores on individual scales were associated with one another and with the overall classification of improvement was tested with ANOVA and Pearson correlations. Analysis of covariance (ANCOVA) was used to test whether overall classification also represented change on individual scales, while controlling for initial value (a residual change comparison, see Chapter 2). Finally, ANOVA, ANCOVA, chi-square analyses, partial correlations, and logistic regression were used to judge whether demographic variables, treatment parameters, or initial assessments are predictive of membership in improver versus non-improver categories or in residual change of individual scales.

3.2.5 Description of treatment

Treatments in this study were delivered by qualified psychoanalysts, and members of the British Psychoanalytical Society, who were trained in the Contemporary Freudian tradition and strongly influenced by the work of Joseph and Anne-Marie Sandler (Sandler & Dreher, 1996). The treatments were strongly focused on transference interpretations. Payment for sessions was set on a sliding scale based on the capacity of the patient to pay, and varied between 1 and 15 pounds per session. In the initial phase of the project, between August 1990 and September 1993, 14 subjects began intensive five times a week psychoanalysis with members of the group (the fifteenth intensive subject studied, subject U, began in January 1995). This was the only treatment option considered by the group at this point and once the subject was accepted into the study they were assigned

to this treatment regardless of presenting symptoms or diagnosis. Starting in September 1994, new subjects (except for U) were accepted only into non-intensive once weekly analysis. No change was made in the criteria for acceptance into the study at this point and no other treatment options were considered.

For the 14 intensive subjects included in the study, length of treatment varied between 1 and 8 years, with a mean of 4.6 years (SD=2.5) and a median of 4.75 years. Number of assessments ranged between 1 and 5 with a mean of 3.57 (SD=1.3) and a median of 3.5. For the 13 of these subjects with at least two assessments, time between first and last assessment varied between 1.5 and 8 years with a mean of 4.1 years (SD=2.0) and a median of 4 years. For the 11 non-intensive subjects included in the study, length of treatment varied between 2 months and 3.5 years, with a mean of 1.5 years (SD=1.1) and a median of 1 year. Number of assessments ranged between 1 and 3 with a mean and median of 2 (SD=0.9). For the eight of these subjects with at least two assessments, time between first and last assessment varied between 6 months and 3 years with a mean of 1.6 years (SD=0.9) and a median of 1.5 years.

Four of the intensive psychotherapy subjects in the study had a change in the frequency of their sessions during the course of the treatment. Subject C dropped from five-times-a-week sessions to three-times-a-week in November 1996, and once-a-week in April 1999. These changes were in response to her refusing to attend at the frequency that she originally agreed to and her analyst recommended. Subjects D, E, and O decreased the frequency of their sessions as part of an agreement with their analysts to move towards termination. Subject D dropped to three-times-a-week in September 1998, but subsequently returned to five-times-a-week in Jan 1999, and then dropped to four-times-a-week in May 1999. Subject E decreased to once-a-week in September 1997. Subject O dropped to four-times-a-week in April 1998 and then once-a-week in Spring 1999.

Attendance of the patients in psychotherapy and psychoanalysis was recorded via monthly attendance sheets filled out by the analysts. From the beginning of treatment through termination or the end of May 1997 (whichever came first) each day was identified as a session, a holiday (a routine and prearranged time of absence, determined by the analyst), an analyst absence (a non-routine time which the analyst could not make a session, because of personal reasons), a patient cancellation, or a patient no-show (with no warning to the therapist that they would not be attending). In addition, several analysts periodically did not turn in attendance sheets and were later unable to report whether sessions had taken place during that period of time. Number of total possible sessions (five-times-a-week for intensive patients and once-a-week for non-intensive patients, excluding holidays) during the treatment, percentage of these days classified as sessions, analyst absence, patient cancellation, patient no shows, and missing data, and estimated number of attended sessions are reported in Appendix 3.1. Summary statistics for intensive and non-intensive subjects are reported in Table 3.4.

Mean (SD)	Possible Sessions	% Sessions	% Analyst absence	% Patient cancellation	% Patient no-show	% Missing Data	Estimated sessions attended
Intensive (n=14)	794 (397)	67 (16)	3.3 (2.9)	13 (16)	8.6 (15)	7.4 (8.4)	582 (344)
Non-intensive (n=10)	51 (35)	65 (19)	1.9 (2.9)	18 (22)	15 (16)	0 (0)	34 (23)

Table 3.4. Summary of attendance data for intensive and non-intensive subjects.

Pearson correlation, ANOVA, and chi-square analyses were used to detect associations between treatment parameters (treatment frequency, treatment length, time between first and last assessment, number of assessments, estimated percentage of sessions attended, and estimated number of sessions attended) and demographic variables (Appendix 3.1). Of 24 statistical tests conducted, four significant results were obtained using an α of 0.05. Subjects in intensive therapy were of a higher socio-

economic class than non-intensive subjects ($F(1,19)=-6.7, p<.05$). Male subjects underwent a higher average number of assessments ($F(1,23)=5.4, p<.05$), and there was a positive association between initial IQ and number of assessments ($r(n=14)=.61, p<.05$). There was also a positive association between subject age at start of treatment and the estimated percentage of sessions attended ($r(n=23)=.43, p<.05$).

3.3 Results

3.3.1 Summary of assessment measures

Original data on all assessment measures is given in Appendix 3.2. Mean values and standard deviations for BDI total, STAI-State, STAI-Trait, SAS-M total, number of SADS-L current diagnoses, and number of SCID-II current diagnoses at initial and termination assessments are reported for intensive and non-intensive groups in Table 3.5. Mean values and standard deviations for EPQ (subscales E, P, N, and L), estimated full-scale IQ, BPRS, and DIB at intake only are reported for intensive and non-intensive groups in Table 3.6. The number of subjects in the clinical range for measures with initial and termination assessments are given in Table 3.7 and for measures with initial assessments only in Table 3.8. Clinical cutoffs were not appropriate for the EPQ L scale, IQ, and BPRS and are therefore not included in Table 3.8.

DSM-III-R Axis I (SADS-L) and Axis II (SCID-II) diagnoses were tabulated for initial, termination, and follow-up assessments for each of the intensive and non-intensive subjects. The frequencies of each diagnosis at initial, termination, and follow-up for intensive and non-intensive subjects is reported in Table 3.9. Mood disorder diagnoses were divided between major depression, dysthymia, and bipolar disorder. It was found that of the two psychiatrist interviewers, one gave several diagnoses of major depression and no diagnoses of dysthymia, while the other gave several diagnoses of dysthymia and only one diagnosis of major depression. Therefore, the two diagnoses

were merged. Only three of the 27 total diagnoses of mood disorders were of bipolar disorder. The seven diagnoses of anxiety disorder were divided between overlapping categories of generalised anxiety disorder (n=3), obsessive compulsive disorder (n=1), and panic disorder (n=5). The four eating disorder diagnoses were all for bulimia, one of which was comorbid with anorexia. Substance abuse diagnoses included both alcohol and hard drug use. Diagnosis of suicidality included patients with severe suicidal ideation who had already made suicide attempts and those with severe suicidal ideation who had not. DSM-III-R Axis II cluster A diagnoses (total n=6) included paranoid (n=5) and schizotypal (n=1) personality disorders. Cluster B diagnoses (total n=12) included overlapping categories of borderline (n=10), narcissistic (n=5), histrionic (n=1), and antisocial (n=1) personality disorders. Cluster C diagnoses (total n=21) included overlapping categories of avoidant (n=11), self-defeating (n=6), dependent (n=2), obsessive-compulsive (n=1), and passive-aggressive (n=3) personality disorders.

Mean (SD) (n)	<i>SCL-90-R</i>	<i>BDI</i>	<i>STAI-S</i>	<i>STAI-T</i>	<i>SAS-M</i>	# <i>Axis I</i> <i>Diags</i>	# <i>Axis II</i> <i>Diags</i>
Intensive							
Initial		20 (10) (n=14)	52 (15) (n=14)	56 (11) (n=14)		1.4 (1.3) (n=14)	0.9 (0.9) (n=14)
Follow-along	1.1 (0.6) (n=9)	15 (11) (n=9)	49 (12) (n=9)	51 (10) (n=9)	2.2 (0.3) (n=9)	1.1 (0.9) (n=9)	1.0 (0.9) (n=9)
Termination	0.6 (0.6) (n=12)	7 (10) (n=12)	39 (10) (n=11)	44 (12) (n=11)	1.9 (0.4) (n=12)	0.4 (1.0) (n=9)	0.5 (0.7) (n=11)
Follow-up	0.8 (0.8) (n=9)	8 (8) (n=9)	44 (18) (n=7)	48 (13) (n=8)	2.0 (0.6) (n=9)	0.7 (1.0) (n=7)	0.9 (1.4) (n=8)
Non-Intensive							
Initial	1.6 (1.0) (n=9)	21 (15) (n=11)	57 (8) (n=11)	59 (10) (n=11)	2.7 (0.7) (n=9)	1.3 (1.1) (n=11)	1.6 (1.0) (n=9)
Termination	1.1 (0.6) (n=7)	14 (6) (n=7)	52 (9) (n=6)	54 (5) (n=7)	2.5 (0.4) (n=7)	0.8 (0.8) (n=6)	0.7 (0.5) (n=6)
Follow-up	0.9 (0.5) (n=5)	17 (11) (n=5)	44 (10) (n=5)	49 (9) (n=5)	2.3 (0.5) (n=5)	0.7 (0.6) (n=3)	0.5 (0.7) (n=2)

Table 3.5. Summary of assessment scales at initial, follow-along, termination, and follow-up for intensive and non-intensive patients.

Mean (SD)	<i>E</i>	<i>P</i>	<i>N</i>	<i>L</i>	<i>IQ</i>	<i>BPRS</i>	<i>DIB</i>
Intensive Initial	10 (7) (n=13)	9 (3) (n=13)	18 (6) (n=13)	6 (5) (n=13)	121 (3) (n=13)	18 (8) (n=12)	5 (2) (n=14)
Non-Intensive Initial	14 (3) (n=2)	15 (11) (n=2)	21 (2) (n=2)	10 (6) (n=2)	125 (n=1)	13 (12) (n=4)	4 (1) (n=4)

Table 3.6. Summary of initial-only assessment scales for intensive and non-intensive patients.

# Clinical (# Non-clinical)	<i>SCL-90-R</i>	<i>BDI</i>	<i>STAI-S</i>	<i>STAI-T</i>	<i>SAS-M</i>	# <i>Axis I</i> <i>Diags</i>	# <i>Axis II</i> <i>Diags</i>
Intensive							
Initial		9 (5)	10 (4)	13 (1)		9 (5)	8 (6)
Follow-along	7 (2)	3 (6)	6 (3)	8 (1)	7 (2)	6 (3)	5 (4)
Termination	6 (6)	1 (11)	4 (7)	4 (7)	5 (7)	2 (7)	5 (6)
Follow-up	3 (6)	1 (8)	3 (4)	5 (3)	4 (5)	3 (4)	4 (4)
Non-Intensive							
Initial	7 (2)	5 (6)	11 (0)	10 (1)	8 (1)	9 (2)	9 (0)
Termination	6 (1)	3 (4)	6 (0)	7 (0)	7 (0)	4 (2)	4 (2)
Follow-up	3 (2)	3 (2)	3 (2)	4 (1)	4 (1)	2 (1)	1 (2)

Table 3.7. Numbers of intensive and non-intensive subjects in the clinical range at initial, termination, and follow-up (number of subjects not in clinical range in parentheses).

# Clinical (# Non-clinical)	<i>E</i>	<i>P</i>	<i>N</i>	<i>DIB</i>
Intensive Initial	5 (8)	13 (0)	12 (1)	4 (10)
Non-Intensive Initial	1 (1)	2 (0)	2 (0)	0 (4)

Table 3.8. Numbers of intensive and non-intensive subjects in the clinical range at initial-only (number of subjects not in clinical range in parentheses).

	Axis I					Axis II		
	Mood disorder	Anxiety disorder	Eating disorder	Substance abuse	Suicidal	Cluster A	Cluster B	Cluster C
Intensive								
Initial	8 (n=14)	3	2	1	2	1 (n=14)	5	5
Follow-along	9 (n=9)	2	2	1	2	1 (n=9)	5	2
Termination	1 (n=9)	0	1	1	1	1 (n=11)	2	1
Follow-up	1 (n=7)	1	0	0	1	1 (n=8)	1	1
Non-Intensive								
Initial	9 (n=11)	2	1	0	1	3 (n=9)	2	7
Termination	4 (n=6)	0	1	0	0	3 (n=6)	0	1
Follow-up	1 (n=3)	0	1	0	0	0 (n=2)	0	1

Table 3.9. Numbers of intensive and non-intensive subjects with DSM-III-R Axis I and Axis II diagnoses at initial, termination, and follow-up.

3.3.2 Demographics versus initial assessment measures

Pearson correlations, ANOVA F-tests, and chi-squares (with Yates correction applied) comparing demographic variables and initial assessment measures are reported in Appendix 3.3. Five of the 68 associations measured were statistically significant using an α of 0.05. Female subjects were more likely to be diagnosed at initial assessment with a mood disorder ($\chi^2(1, n=25) = 10.0, p < .01$), had a higher mean number of Axis I diagnoses ($F(1, 23) = 6.6, p < .05$), and had a higher total score on the DIB ($F(1, 16) = 5.7, p < .05$) than men. Gender was not, however, significantly related to score on the BDI, number of Axis II diagnoses, individual clusters of personality disorders, or any other initial assessment measures. Socio-economic status was associated with average SCL-90 and SAS scores. Subjects with lower SES scored higher on each measure (SCL-90: $r(n=18) = .53, p < .05$; SAS: $r(n=18) = .61, p < .01$). SES was not associated with initial assessment measures. Age of subject at start of treatment and estimated full scale IQ were not associated with any of the initial assessment measures.

3.3.3 Treatment parameters versus initial assessment measures

Statistical associations between treatment parameters and initial assessments are reported in Appendix 3.3. Of the 102 statistical tests applied, only one is significant using an α of 0.05. Number of Axis I diagnoses at initial assessment was found to be negatively associated with the estimated percentage of available therapy sessions attended by the subject ($r(n=24)=-.43, p<.05$).

3.3.4 Concurrent validity of assessment measures

Associations between all initial-only and initial-termination assessment measures are reported using Pearson correlations, ANOVAs, and Yates corrected chi-square analyses, where appropriate, in Appendix 3.3. Significant positive associations ($r=.67$ to $.86$) were found between every pairing of the five self-report scales (SCL-90-R, BDI, STAI-State, STAI-Trait, and SAS-M). Number of Axis I diagnoses was positively associated with SCL-90-R, BDI, STAI-S, and STAI-T. Number of Axis II diagnoses was positively associated with SCL-90-R, STAI-T, and SAS-M. Number of Axis I and number of Axis II diagnoses were not significantly associated. A diagnosis of mood disorder was associated with higher SCL-90-R, BDI, and STAI-T scores. A diagnosis of cluster C personality disorder (mostly avoidant personality disorder) was associated with higher SCL-90-R, BDI, STAI-T, and SAS-M scores. There were no significant relationships between diagnoses of mood disorders, or cluster A, B, or C personality disorders.

EPQ extroversion scale (EPQ-E) was negatively associated with STAI-S and STAI-T. EPQ neuroticism scale (EPQ-N) was positively associated with SCL-90-R, BDI, STAI-T, mood disorder, and BPRS. BPRS was also positively associated with STAI-S, STAI-T, number of Axis II diagnoses, and mood disorder. EPQ lying scale (EPQ-L) was positively associated with SCL-90-R. EPQ-E and EPQ-P were positively associated.

3.3.5 *Change on individual measures*

Mean values at initial assessment and termination on the four repeated assessment scales are depicted in Figures 3.1a through 3.1d (n=19). When subjected to ANCOVA, treating initial value as the covariate, all of the scales, except number of Axis II diagnoses, showed statistically significant change. Change from initial assessment to termination are displayed separately for intensive (n=12) and non-intensive subjects (n=7) in Figures 3.2a through 3.2d. Analyses were not performed separately on intensive and non-intensive subjects because of the small number of assessments per time period. Figures 3.3a through 3.3d show mean values at initial assessment, follow-along (itself an average of all follow-along points available for an individual), and termination for each of the scales (n=10). Improvement on each of the scales appears gradual with some change by follow-along and further change by termination. Statistics were not performed on these values due to the small sample size. Finally, mean values at initial assessment, termination, and follow-up (18 months after termination) are shown in Figures 3.4a through 3.4d (n=13). The improvement achieved by termination is maintained on the BDI, STAI-T, and number of Axis I diagnosis scales. On the number of Axis II diagnoses scale there was some deterioration after termination, though the mean was still less than it had been at initial assessment (recall that in the larger sample, the improvement in number of Axis II diagnoses did not reach statistical significance by termination).

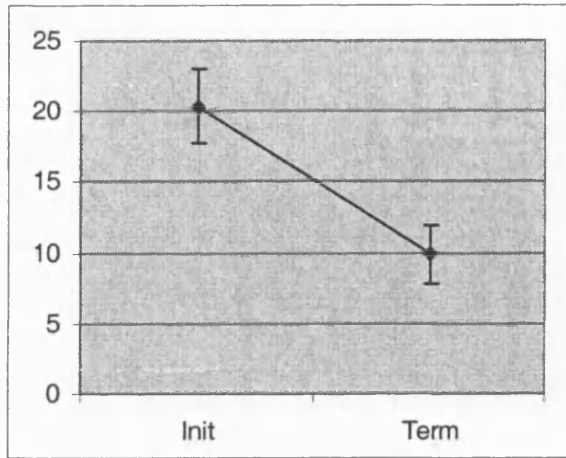


Figure 3.1a. BDI.

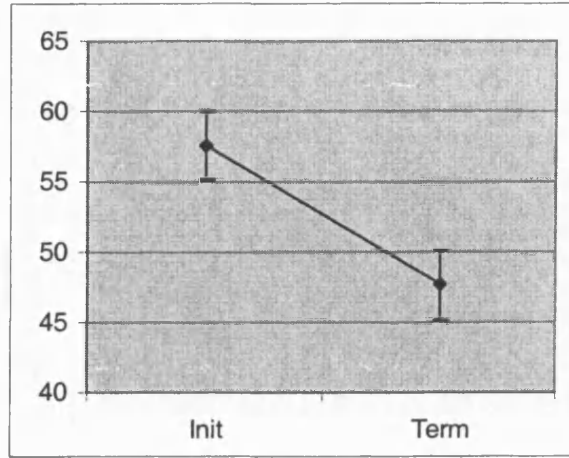


Figure 3.1b. STAI-T.

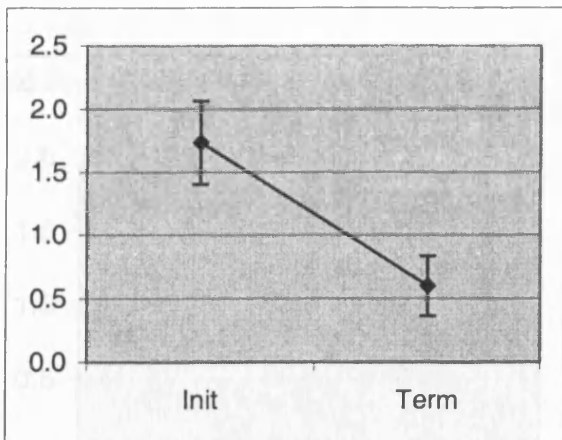


Figure 3.1c. Axis I Diags.

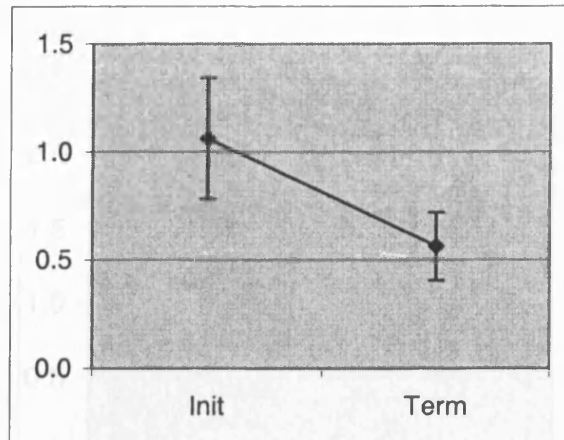


Figure 3.1d. Axis II Diags.

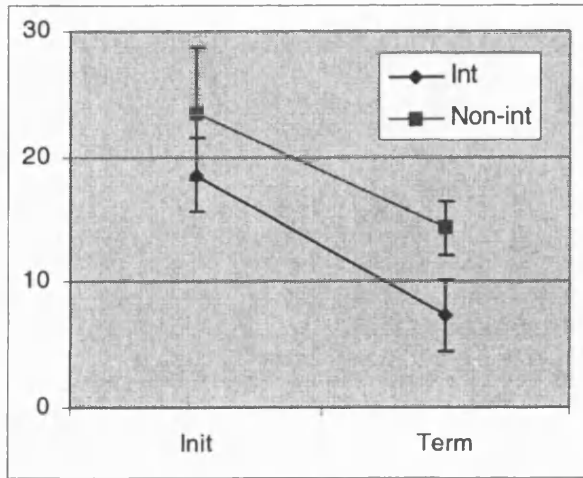


Figure 3.2a. BDI.

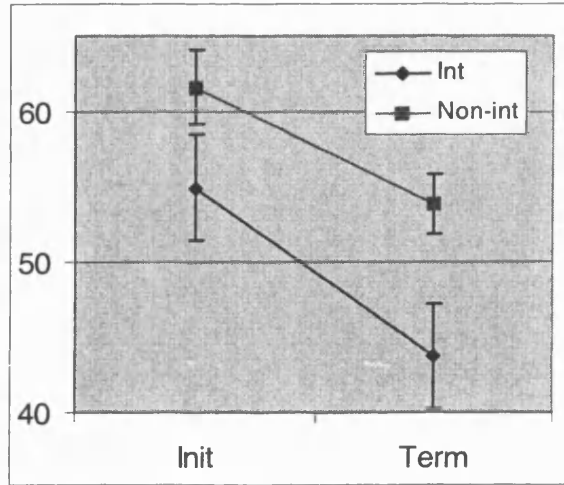


Figure 3.2b. STAI-T.

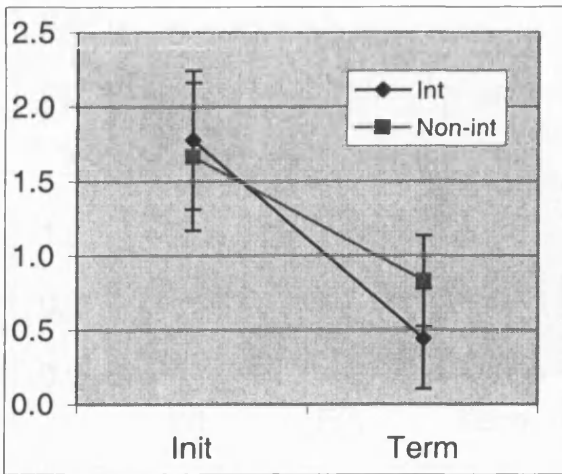


Figure 3.2c. Axis I Diags.

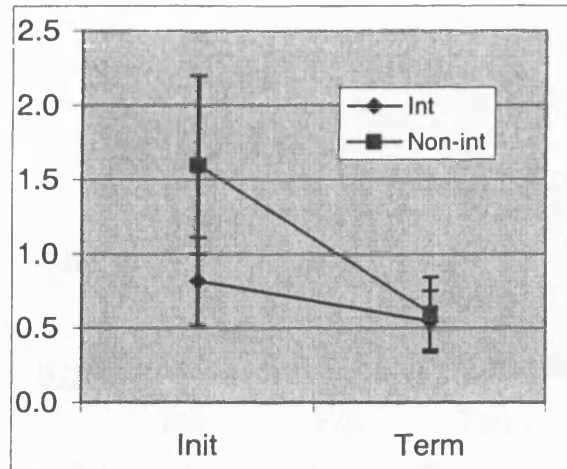


Figure 3.2d. Axis II Diags.

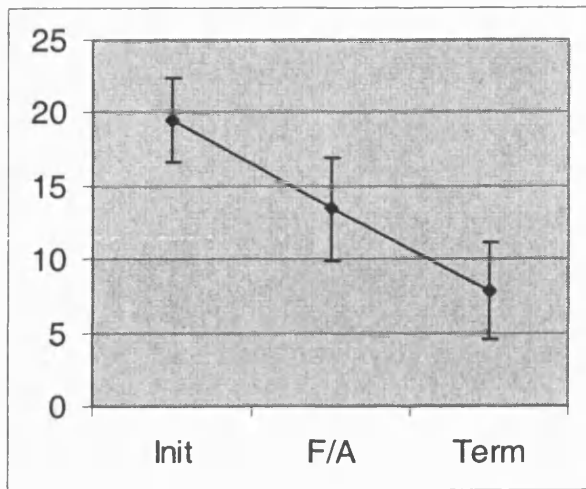


Figure 3.3a. BDI.

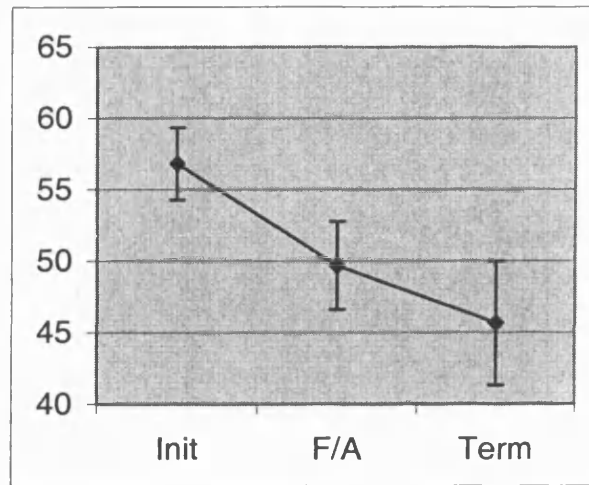


Figure 3.3b. STAI-T.

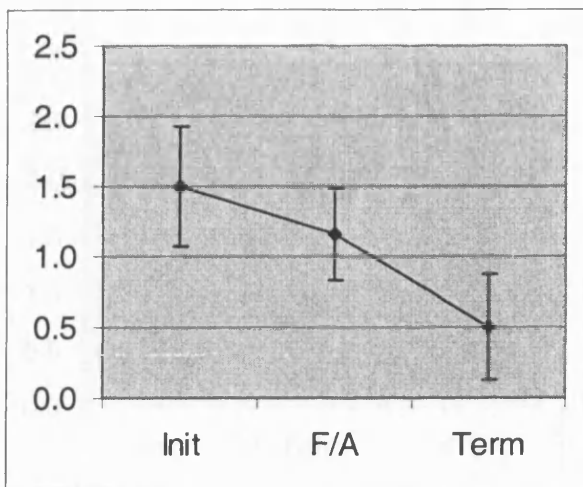


Figure 3.3c. Axis I Diags.

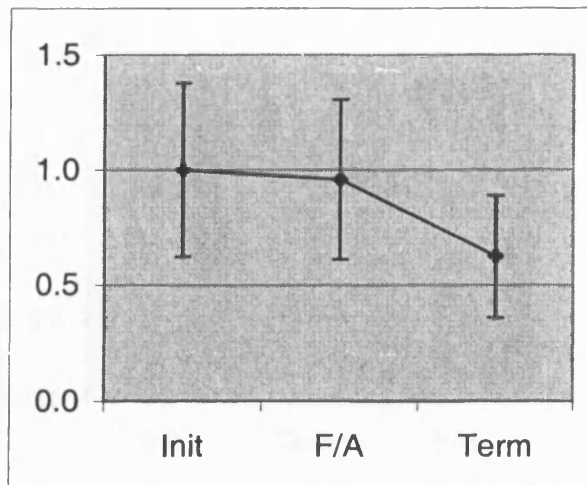


Figure 3.3d. Axis II Diags.

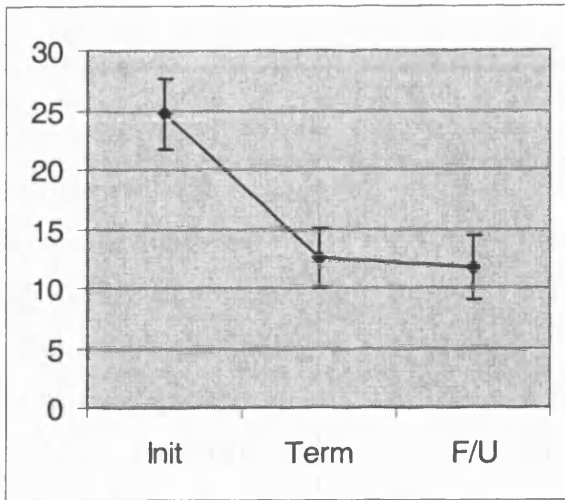


Figure 3.4a. BDI.

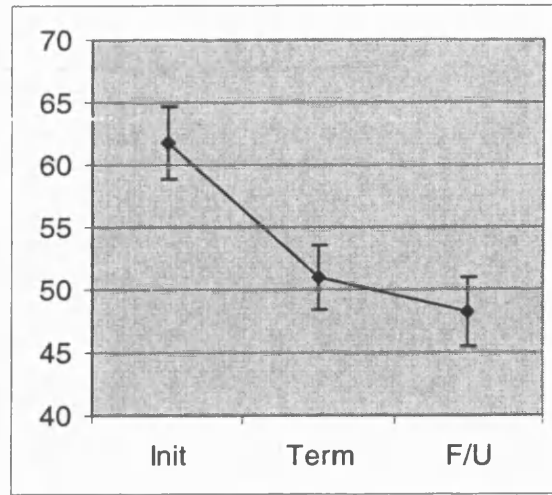


Figure 3.4b. STAI-T.

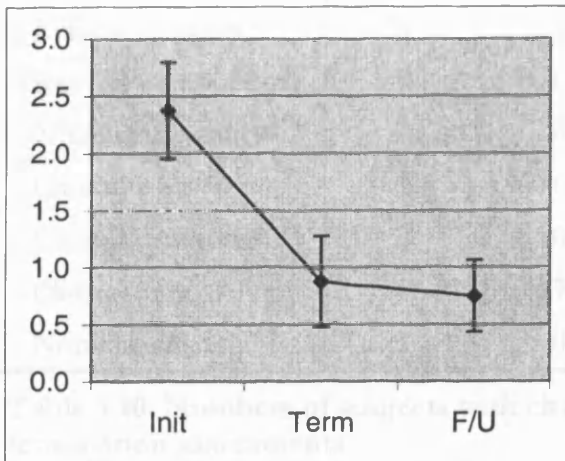


Figure 3.4c. Axis I Diags.

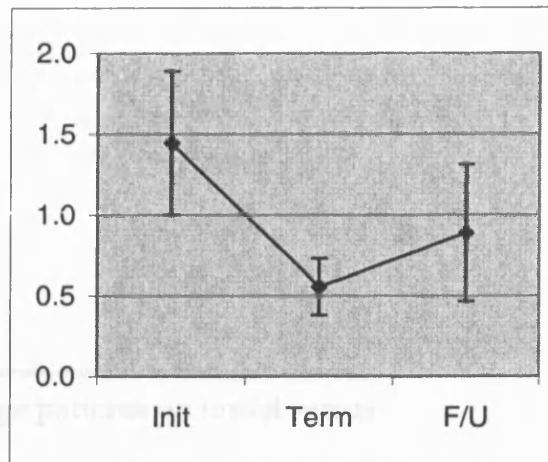


Figure 3.4d. Axis II Diags.

3.3.6 Determining overall change categories

Numbers of subjects with reliable improvement, reliable deterioration, no reliable change, clinical improvement, clinical deterioration, clinical only, and non-clinical only status on initial versus termination measures are reported in Table 3.10. Due to the high correlation, reported above, between STAI-State and STAI-Trait, and the greater suitability of STAI-Trait for measuring change in response to psychotherapy, as indicated in the literature on the STAI, only STAI-T was used for measuring change from initial to termination. Breakdown by subject is reported in Appendix 3.4.

	<i>BDI</i>	<i>STAI-T</i>	<i># Axis I Dx</i>	<i># Axis II Dx</i>
Intensive (n=12)				
Reliable improver	8	5	4	2
Reliable deteriorator	0	0	0	0
No reliable change	4	6	6	9
Clinical improver	6	7	5	3
Clinical deteriorator	0	0	0	1
Clinical only	1	3	2	3
Non-clinical only	5	1	3	4
Non-intensive (n=7)				
Reliable improver	3	1	1	1
Reliable deteriorator	0	0	0	0
No reliable change	4	6	6	6
Clinical improver	3	0	2	2
Clinical deteriorator	2	0	0	0
Clinical only	1	7	4	5
Non-clinical only	1	0	1	0

Table 3.10. Numbers of subjects with change patterns on initial versus termination assessments.

Determination of overall improvement status from the results produced so far had several goals. First, it should capture the diversity of change; subjects who take different paths to improvement and/or record their improvement on different measures should all be considered improvers. Second, for purposes of reliability, it should require significant improvement on at least two of the four scales on which improvement was assessed. Third, it should take into consideration both reliable change, i.e., change of such degree that it is unlikely to have occurred because of instability of the measure, and clinical change, i.e., movement from the clinical to non-clinical groups (or vice versa) which can be important even if the numeric change in the scale is relatively small. In order to

accommodate all these goals, a formula was chosen to calculate a change index for each subject:

$$\text{Change index} = \frac{\begin{aligned} &(\text{Number of scales on which reliable improvement occurred} \\ &+ \text{Number of scales on which clinical improvement occurred}) \\ &- (\text{Number of scales on which reliable deterioration occurred} \\ &+ \text{Number of scales on which clinical deterioration occurred}) \end{aligned}}{\text{Number of scales in the clinical range at initial assessment}}$$

If a subject had a change index of 1 or greater, corresponding to reliable and/or clinical change on at least half of the scales initially in the clinical range, he/she was considered to be an overall “improver.” Since no subjects had change indices of less than 0, all other subjects were designated as “non-improvers.” Of the 12 intensive subjects for whom change information was available, three had indices of 2, two had indices between 1 and 2, five had indices of 1, and two had indices of 0. One of the subjects with an index of 0, did not score in the clinical range on any scales at initial assessment, but as none of his scales had even shown any reliable change, he was classified as a non-improver. Of the seven non-intensive subjects for whom change information was available, two had indices of 1 or greater, three had indices between 0 and 1, and 2 had indices of 0. Thus, ten of 12 intensive subjects and two of seven non-intensive subjects were designated as overall improvers.

3.3.7 Comparison of improvers and non-improvers

Associations of demographic variables, treatment parameters, and initial assessment measures with improvement status are reported in Appendix 3.5. No demographic variables were significant predictors of improvement. Treatment intensity and number of sessions attended predicted improvement, but treatment length, number of assessments, and percentage of sessions attended did not. Of the 15 initial assessments tested (BDI, STAI-T, number of Axis I diagnoses, number of Axis II diagnoses, mood

disorder, suicidality, substance abuse, eating disorder, anxiety disorder, BPRS, DIB, and four EPQ scales) none predicted improvement significantly.

3.4 Discussion

3.4.1 *Questions raised in the introduction*

The summary tables, statistical tests, and figures presented above provide a general picture of the sample, the treatment provided, and the change in symptoms and diagnoses during treatment. The specific questions raised in the introduction to this chapter about the inter-relationships of demographic variables, initial assessments, treatment parameters, and treatment response are each addressed by the results and will be discussed below. However it is clear throughout that the ability to answer any of these questions authoritatively is limited by the small sample size and the missing data at various time points and on various measures. In particular, small cell size prevents a statistically useful analysis of the differential effect of intensive and non-intensive treatment on patients with different demographics or initial diagnoses. For example, even the interesting finding of a significantly higher rate of improvement in the intensive treatment group has little implication for a causal hypothesis because of potential confounders (treatment length, percentage of sessions attended, SES, and initial severity of illness). Interestingly, while all four significantly differed between intensive and non-intensive groups, only treatment length and number of sessions attended significantly distinguished between the improver and non-improver groups.

This brings to the forefront two facts of this study which should be kept in mind while reading the early discussion and will be returned to later in this chapter: (1) Any increase in the number or application frequency of a study's measures (both of which are unusually high in this study as compared to the psychotherapy and psychoanalytic literature) brings with it the practical consequence of fewer subjects and more missing

data. A study, such as this one, whose aim is more to understand how subjects improve and to propose new measures of structural change (rather than to carefully count the rate of improvement in different subsamples), has obeyed this principle in seeking to optimize the data set for the questions proposed. (2) Given that all studies of psychoanalysis and psychodynamic psychotherapy are to some extent naturalistic (i.e., within the gross constraint of the type of treatment assigned, the process is allowed to unfold according to the natural negotiation of patient and therapist), a study that looks at process cannot avoid a large number of confounding variables. The only way to adequately account for these confounders is to collect broader data, which itself leads to a practical reduction in sample size. Therefore, thorough studies of process have much greater difficulty in providing conclusive answers to statistical questions requiring large sample sizes and homogeneity of subsamples. In summary, there is a balance between the extent to which a study is good at (a) introducing new methodologies and studying structural change and process, and (b) answering demographic, treatment parameter, and outcome questions. This study was designed to focus primarily on (a), but this chapter addresses (b).

3.4.2 Demographics and initial assessments

Associations between demographic variables and initial assessment measures were in accordance with known patterns of psychopathology. According to the DSM-IV major depressive disorder is twice as prevalent in women as in men (1994). As depression was the most common Axis I disorder reported, it is understandable that female subjects also had a higher mean number of Axis I diagnoses. Gender is not, however, associated with BDI (Beck & Steer, 1993), reminding us that an elevated BDI score and a diagnosis of depression are distinct phenomena. In most populations, borderline personality disorder (BPD) is predominantly found in women (75% of cases; DSM-IV, 1994) explaining the association with DIB. Interestingly, DIB and SCID-II

diagnosis of BPD (or Cluster B personality disorders) were not related, nor were gender and SCID-II BPD diagnosis. These may indicate an important difference in the way DIB and SCID-II classify BPD. However, non-significant findings in such a small sample, only half of which had a Cluster B personality disorder diagnosis, are hardly conclusive.

Elevation of SCL-90 and SAS in subjects of lower SES is consistent with the long-standing epidemiological finding that psychiatric distress is more common in persons of lower class (Kaplan, Sadock, & Grebb, 1994). Non-significant findings with age and IQ are not surprising given equivocal literature in this area, but may also be due to the limited range of both variables within our sample. Negative association between number of Axis I disorders and treatment attendance may indicate that poor attendance was a sign of increased psychopathology. It is not clear whether this is related more to a certain type of psychopathology, as is indicated in the literature (Baruch et al., 1998).

High correlations among assessment measures support Lambert's (1994) argument that while it is important to use multiple measures to assess outcome, construct validity of these measures is often poor, and there is considerable overlap in what they are measuring. This overlap is indistinguishable from comorbidity, which is evident from multiple Axis I and Axis II diagnoses at initial assessment in 45% of the sample. Though, such comorbidity is troublesome from the perspective of neatly characterizing the sample, it undoubtedly reflects the true nature of clinical samples and thus supports the generalisability of findings emerging from this study (Richardson, 2001).

3.4.3 Treatment outcome

Significant improvement by termination on three of the four assessment scales when averaged across intensive and non-intensive samples supports the assertion that patients do generally improve on a variety of diagnostic and symptoms scales when undergoing psychoanalysis or psychodynamic psychotherapy. This finding is significant despite the fact that the sample included several early drop-outs.

The lack of a control group prevents the treatment outcome data from telling us anything about whether psychoanalysis and psychotherapy are more effective than no treatment at all and the small sample size prevents us from using this technique to directly compare efficacy of psychoanalysis and psychodynamic psychotherapy. It is worth noting that the only scale that did not show a statistically significant improvement is the number of Axis II diagnoses. This may be related to the inherent instability of this measure (see Methods) or the relative insensitivity to change of a measure that asks questions about general patterns of feeling and behaviour. Most provocatively, this finding may support existing evidence that personality disorders are both less likely to recede spontaneously and less amenable to change via psychotherapy than Axis I disorders (e.g., Berger et al., 2004).

The number of subjects with reliable or clinical change on any of the individual scales is smaller than expected based on reports such as Sandell (2000) who found an increase of 12% to 70% in the number of subjects who were below clinical cutoff on three measures (SCL-90, SAS, and Sense of Coherence). The number of subjects with reliable improvement on a given measure, combining across intensive and non-intensive groups, ranged from 3, for number of Axis II disorders, to 8 for BDI. The number of subjects with clinical improvement was similarly low, ranging from 5 for number of Axis II disorders to 9 for BDI or number of Axis I disorders. This suggests either that improvement in this sample was generally small, or, improvement occurred in such a broad range of client attributes that no one measure was able to capture the effect. This supports our aim in using a composite index of reliable and clinical change across all six measures as a way to sample multiple symptom and diagnostic domains.

The overall rate of improvement, using the composite measures approach, was good (63%) in this sample and is in line with the typical rates reported by other studies of psychoanalysis and psychodynamic psychotherapy (Fonagy et al., 2001). It is naturally

tempting to conclude from the statistically higher rate of improvement in the psychoanalysis group than in the psychotherapy group, that psychoanalysis is a more effective treatment in this population. This is supported by the positive association between number of sessions attended and improvement and the lack of association seen with potential confounding variables (SES, gender, age, IQ, percentage sessions attended, and even length of treatment in weeks). However, as made clear by decades of previous attempts to draw such conclusions, several important confounding factors must be better studied before any solid inferences are drawn. First, in this study, as is common in clinical practice, intensive treatments lasted, on average, significantly longer than non-intensive treatments, and our sample size is too small to distinguish the effect of intensity versus length of treatment. Second, due to difficulties collecting comprehensive follow-up data, improvement status was determined on the basis of a termination assessment, and we do not know if improvements were sustained at follow-up. Previous studies have suggested (Bateman & Fonagy, 1999) that psychoanalytic treatments have longer lasting benefit than non-psychoanalytic treatments, and theories of structural change and psychoanalytic mechanism of action would predict that the more intensive treatment is likely, if anything, to look even better when follow-up assessments are taken into account.

Visual inspection of the trends towards improvement in intensive and non-intensive groups subdivided by assessment measures (see Figures 3.2a through 3.2d) suggest that the advantage of intensive treatment is best captured by the anxiety trait (STAI-T) and depression (BDI) measures, but not by the numbers of Axis I and Axis II diagnoses. Presence or absence of a diagnosis is a fairly crude measure of psychopathology and it is expected that a continuous symptoms measure does a better job at capturing smaller differences in treatment improvement.

This study supports prior psychoanalytic and psychodynamic psychotherapy research in failing to predict those who will respond to treatment from initial diagnostic

assessment measures (Fonagy et al., 2001; Kantrowitz, 1993). Given that DSM-based diagnostic schemes are symptom-based and do not capture the underlying structure of psychopathology, it is not surprising that these are poor predictors of outcome (Garfield, 1998; Richardson, 2001). To date, only one methodology has consistently differentiated those patients who improve in psychoanalysis and psychotherapy on the basis of initial assessments (Blatt & Ford, 1994; Blatt, Quinlan, Plikonis, & Shea, 1995; Blatt & Shahar, 2004). Blatt and colleagues have shown in two retrospective re-analyses of data that patients with more ruminative, self-reflective (i.e., “introjective”) personality configurations improve more in psychoanalysis, while patients with dependent, unflective, more affectively labile (i.e., “anaclitic”) configurations do better in supportive-expressive psychotherapy. Such an assessment strategy was not available at the time of this study, but could conceivably be applied through a re-analysis of existing measures.

The lack of significant differences between intensive and non-intensive groups on percentage attendance, demographics, and initial assessments lends support to the comparability of these subsamples, though the small numbers of subjects limits the power for accurately measuring real differences which may have affected the observed difference in improvement rates. Finally, the fact that the therapists in this study were all trained psychoanalysts, who believe in psychoanalysis as a treatment for personality-disordered young adults, raises the possibility of a treatment allegiance effect, as described by Luborsky and colleagues (1999). They reviewed several previous studies and one of their own demonstrating a strong correlation (as high as $r=0.85$) between the allegiance of a team of researchers to a particular psychotherapeutic technique, as measured by a self-report questionnaire, and the tendency for them to find evidence that this technique is more efficacious than others. Although, the absolute size of the effect predicted by researcher allegiance is still quite small (Lambert, 1999) and no mechanism by which researchers are influencing their results has been identified, it is not surprising

in a field filled with difficult to observe findings and uncertain methodologies that researcher allegiance plays a role in affecting the outcome.

3.4.4 Follow-along and follow-up analyses

Due to missing assessments at follow-along and follow-up time points, it was not possible to differentiate intensive and non-intensive treatment subjects using this data and statistical analyses were not conducted. However, Figures 3.3a through 3.3d and 3.4a through 3.4d reveal some interesting patterns, suggesting overall trends in the follow-along and follow-up data. The generalizability of these trends is supported by the lack of significant difference between subgroups on demographic, treatment parameter, and initial assessment data as demonstrated in the group analyses reported above. Looking only at the 10 subjects for whom initial, follow-along, and termination data were available, the improvement in BDI scores, STAI-T scores, and number of Axis I diagnoses appears gradual and steady, supporting the hypothesis that there is a linear relationship between either treatment dose or time and symptom improvement. Number of Axis II diagnoses do not appear to change at all by follow-along assessment, supporting the view of Axis II pathology as more stable and less amenable to change. Axis II diagnoses appear to begin to retreat by termination, raising the important question of how these diagnoses fare at follow-up. In the separate analyses of 13 subjects for whom follow-up data were available (only 6 subjects overlapped between the follow-along and follow-up samples), improvement was maintained on BDI, STAI-T, and number of Axis I diagnoses. Interestingly, when looking only at these 13 subjects, there appeared to be a significant improvement in number of Axis II diagnoses from initial to termination assessment, but this improvement vanished at follow-up. A larger and more complete data set would be needed to ascertain whether this represents an artifact of the data, or whether psychoanalysis and/or psychodynamic psychotherapy achieve some improvement in Axis II pathology that is not sustained after termination.

3.3.3 Problems with study design and statistical analysis

The most significant flaws in this study are the small sample size and the inconsistency of data collected at follow-along (15 missing assessments), termination (6 missing assessments), and follow-up (12 missing assessments). As suggested above, this was a consequence of the fact that most of the attention in this study was focused on an intense look at a few subjects very frequently, detracting from the possibility of following a larger number of subjects or each subject with exquisite attention. We believe that this raises the broader point, that contrast-group methodology, which is often held up as the gold-standard in psychotherapy research is not always the optimal means for empirically answering a question. Specifically, when the task is to understand a breadth of measures, structural change, and process, it is worthwhile to move the focus from contrasting large groups to studying small numbers of subjects well.

The current standard for handling missing data in treatment outcome research (particularly in pharmacologic research) is to employ the “last observation carried forward” method. In this system, missing data at the chosen endpoint (such as termination or follow-up) is replaced by the last available measurement. Unfortunately, for the six subjects left out of the data analysis due to missing data (two intensive and four non-intensive), there is no rational way of applying this methodology. Five of these subjects have no assessment after their initial battery and one has only a follow-up assessment 2 years after termination. In all six cases, carrying the last observation forward to termination would result in assigning them to the non-improver group. Given the large number of such subjects relative to the overall sample size this would be too conservative an approach and potentially mask interesting findings from the other 19 subjects. Of course, the current results are potentially biased as they leave out those subjects who may not have continued with the study for reasons that are important to the efficacy of the treatment. However, no significant differences were observed between

the six missing subjects and the remaining 19 on the basis of an analysis of demographics and initial assessments, and it was judged that this bias was preferable to the alternative that made use of the “last observation carried forward technique.” For similar reasons, this technique was not used to analyse the follow-up data. Last observation carried forward (from termination to follow-up) would have been used for 13 of the 25 subjects and surely obscured any interesting findings regarding follow-up. Instead, no statistical analysis of follow-up data was performed.

This study would have benefited from more consistent collection and statistical analysis of certain characteristics of the patients and treatment. Attempts to characterize the nature of the treatment termination were unsuccessful and no statistical analyses were performed to gauge the relationship between termination and outcome. A future study would benefit from a comparison between therapies ending in mutual versus patient-only terminations. It would also have been helpful to better describe certain non-diagnostic aspects of the patient’s current and past experiences and behavior, including physical and sexual abuse, suicidality and parasuicidality, substance use, and previous psychiatric and psychotherapeutic treatments.

Disadvantages of RCTs

Although RCTs are widely accepted as a scientific standard, their feasibility and applicability to psychotherapy research are limited for a number of reasons (Goldfried & Wolfe, 1998; Jadad, 1998; Richardson, 2001; Shadish et al., 1997; Warren & Norton, 2004). Evidence for the severity of these limitations is that the criteria for ESTs suggested by the American Psychological Association (1995; Chambless et al., 1996), specify “good group design,” but not specifically an RCT. The added requirement of randomization can make a study more complex, expensive, and time-consuming (Goldfried et al., 1990; Kazdin, 1986). In addition, randomization alone does not address the important problems of subject nonuniformity, treatment nonuniformity, difficulty of

finding adequate control groups, and nongeneralizability of findings (Fonagy, 2001b; Gabbard et al., 2002). The criteria for “good group design,” may be easier to meet in an RCT, but are no means satisfied automatically and, for practical reasons, may be better met by a different type of investigation.

One methodological problems of particular significance in RCTs is that of attrition (Fonagy, 2001b; Kazdin, 1994). Although an RCT controls initially for group differences by randomly assigning patients to treatments, there is no ethical means by which patients can be prevented from leaving treatments early. Thus, at assessment points during or after therapy, the groups may become biased, regardless of how well they were randomized. Attrition can interfere with the results of an experiment in four ways: (1) subjects who drop out early may differ from those who remain in the study, limiting the generalizability of the results, (2) the characteristics of subjects who drop out from different groups may differ significantly, undoing the effects of randomization, (3) the number of subjects who drop out may vary significantly among groups, confounding the first two problems, and (4) so many cases may drop out that valid conclusions about the treatment are not possible. These sources of error are particularly problematic in studies which evaluate long treatments or depend on long-term follow-up for significant results. Statistical methods have been developed to reduce the impact of attrition on outcome results, including use of the last available data-point to estimate bias (Flick, 1988; Little & Rubin, 1987) and “intention to treat” analyses based on all the subjects who began treatment, regardless of whether they dropped out (Elkin et al., 1989). While the problem of attrition bias may be somewhat easier to isolate in RCTs, it reminds us that good study design is a wider concept.

Disadvantages of treatment package design and dismantling, constructive, and parametric strategies

The use of no-treatment, waiting list, or placebo controls, inherent to the treatment package approach, has been questioned on both methodological and ethical grounds (Fonagy, 2001a, 2001b; Goldfried et al., 1990; Horvath, 1988; Parloff, 1986). According to the standards set forth by the medical model of research, a treatment is shown to be effective if and only if patients undergoing that treatment improve significantly more than a similar group of patients who are treated in exactly the same way, *except for not receiving the active ingredient of the therapy being tested*. Parloff and Horvath (Horvath, 1988; Parloff, 1986) point out that due to lack of standardisation of placebos and psychotherapies and the difficulty in enforcing or monitoring such standards, it is difficult to establish placebos that are similar to therapy but lack the active treatment components. Even if through a sophisticated theory of psychotherapeutic mechanisms of change and manualization such a placebo were designed, it is likely that it would be more like an alternative type of therapy than a non-treatment. In addition, unlike with a placebo pill, it is meaningless to attempt to blind patient and therapist to the difference between treatment and placebo. Due to therapist allegiance effects (Luborsky et al., 1999) and difficulty in finding a placebo that patients will believe is as effective as the experimental treatment, it is difficult to conclude that confounding causes of a placebo's inferiority have ever been ruled out (Parloff, 1986).

The very premise of dismantling, constructive, and parametric strategies relies on the assumption that we are able to theoretically and practically parse out the elements of a treatment into components which we can control and test. At the very least, such an undertaking relies on a theory for what these elements are and a methodology for measuring how much they are present in a given therapy. Kiesler (1966; 1995) has argued that despite advances in a theory of therapeutic action, we are still awaiting further

process studies to give us the data we need to accurately parse therapies into their component parts. It has been well established that therapists employ techniques from prototypes other than their own and that there may be as much variation of therapist techniques within a single treatment group as there is between two supposedly different treatment groups. In studying psychotherapy transcripts of cognitive and psychodynamic short-term therapies for depression, Ablon and Jones (1998) found that outcome was associated with the extent to which treatment in *either group* matched the empirical prototype of psychodynamic psychotherapy. In a second study, analyzing transcripts from the NIMH Collaborative Study for the Treatment of Depression (Elkin et al., 1989), Ablon and Jones (1999) adherence to the cognitive therapy prototype was most predictive of change. In both, supposedly distinct therapeutic methods showed surprising similarities in the range of techniques used.

3.3.4 Implications for future chapters and future research

Importance of process research

Process research is widely recognized as a necessary step if we are to better understand the mechanisms by which different kinds of psychotherapy lead to change in patients with a variety of disorders under a range of circumstances (Blatt, 2001; Greenberg, 1991; Imber, 1992; Kazdin, 1986; Strupp, 1986; Westen & Morrison, 2001). Paul (1967) notes that the “flight to process” in psychotherapy research is in part due to the fact that it circumvents sociological difficulties in evaluating which kinds of therapy, which therapists, or which patients have the best outcome.

Although process research does not directly answer the questions posed by the EST literature, it is undoubtedly a part of evaluating therapies for this purpose. Studies designed with PTO and ATI principles depend on process work for their conceptual foundation. Researchers disagree somewhat in the optimal order in which comparative therapy and process strategies should be employed. Gabbard, Gunderson, and Fonagy

(2002) and Paul (1967) suggest beginning with large scale randomized controlled trials to establish which treatments are optimal for which conditions, and then investigating these patterns further by pursuing more careful outcome work. Other researchers (Greenberg, 1991; Kächele, 1992) advocate using process studies to first identify relevant variables and interactions, thus making it easier to design large scale studies that demonstrate these findings.

Unfortunately, limited resources and the time and money consuming nature of process work limits the extent to which process measures can be collected consistently in large scale studies. This is a serious limitation on both studies, however, because process and outcome measures supplement one another, and one cannot be truly interpreted without the other (Kiesler, 1966; Stiles & Shapiro, 1989, 1994). To be properly analysed process measures require frequent time points and sophisticated means for checking reliability and external validity. Finally, questions have been raised whether, due to their demanding methodology, the implementation of process measures results in added patient and therapist selection bias and limits the generalisability of such research (Roth & Fonagy, in press).

Advantages of single-case and qualitative research

Methods of quantitative data collection within single cases have been progressing at a rapid rate and include self-report, direct observation, and text analytic schemes (Fonagy et al., 2001). Measures for analyzing therapy transcripts either by computer (Bucci, 1997b; Kächele, 1992) or by independent rater (Jones, 2000) are currently being developed. In single case experimental designs, researchers aim to collect process data at multiple stages of the therapeutic process, including at initial, during one or more treatments, and following treatments. Following the dictates of interrupted time-series designs (Cook & Campbell, 1979), researchers often attempt to increase validity by detecting differences between phases of interaction when interventions are being applied

and when they have been withdrawn. Single case design has been applied to all major treatment modalities, though historically experimental designs were more common to behavioural and cognitive-behavioural interventions, while case studies were most common to psychoanalysis (Fonagy, 2001b). Recently, though, psychoanalytic researchers have begun to perform single-case experiments as well (Jones, 2000; Moran & Fonagy, 1987).

The advantages of single case studies are significant. They offer the opportunity for collecting detailed data on the patient-therapist relationship, including the idiosyncrasies and subtleties that would be lost in the wash of aggregation across nonuniform patients and therapists. Single case studies can be incorporated into the clinical practice of private practitioners and often do not involve the expensive and time-consuming overhead of a large scale research project. In a single case experiment, careful manipulation of a single subject is possible and aggregation of repeated measures within the subject gives the study the statistical power that it needs.

Difficulty in generalizing the results of single case research is no longer considered an insurmountable problem. In fact, several researchers have suggested that the only route to generalised rules of psychotherapy is to begin with single cases (Bachrach, 1993; Strupp, 2001). "To find out what people do in general, we must first discover what each person does in particular, then determine what, if anything, these particulars have in common" (Thorngate, 1986). Aggregation across subjects before the individual processes are sufficiently understood can lead to distortion of findings and loss of the "fine grain." (Hilliard, 1993). It has even been suggested that generalizability is overrated, as one is sometimes more interested in what is possible than what is common (Hilliard, 1993). To overcome within subject sources of error, the subject can act as his own control and effects of the intervention can be studied by comparing different conditions presented to the same subject over time (Barlow & Hersen, 1973; Kazdin, 1982). To further

generalizability, single case studies can be replicated directly (i.e., in similar patients), or systematically (i.e., in different subjects, so as to show how findings differ in predictable ways based on subject variation). Hilliard (1993) suggests that lack of both types of replication has been the greatest weakness in the field of single-case research.

The most significant disadvantage of single-case research is limited generalizability, particularly when a patient has been carefully selected, thus supplying a plausible hypothesis for why lessons from this case do not apply to others (Fonagy, 2001b; Kazdin, 1994). The absence of between subject variability eliminates the possibility of studying macroscopic client-treatment interactions and post-hoc analyses of differential responsiveness to treatment. Experimental techniques for demonstrating internal validity are limited by the extent to which the treatment effect is immediate, reversible, and specific to the area of behaviour being measured (Hilliard, 1993; Kazdin, 1982). Experimentally useful manipulations such as withdrawing treatment are not always compatible with clinical practice. Aggregating data within a subject has the same dangers as aggregating across subjects, for example, aggregating data from different phases of therapy may obscure or distort important patterns. Finally, the wide range of single case methodology, particularly towards the case study end of the continuum, makes this strategy susceptible to poor methodological practices.

Qualitative or conceptual research design is a loosely constructed methodology, originating in sociology, for studying subjective processes, usually using narrative analysis for hypothesis generation and testing. Some investigators, particularly in psychoanalysis, have argued that given the intensively subjective nature of the therapeutic encounter, it is appropriate to use the narrative reports of a therapeutic encounter by the therapist as a basic source of data (Dreher, 2000; 1993; Tuckett, 1994). "Ours is interpretive work and ... must include the perspectives and voices of the people whom we study" (Strauss & Corbin, 1994). Grounded theory methodology, developed by Glaser and Strauss (Strauss

& Corbin, 1994), is a further refinement of qualitative design emphasizing theory development and verification of hypotheses throughout the course of a qualitative research project. Like other modes of qualitative research, it uses interviews, field observations, documents, and videotapes as its sources of data, and can incorporate qualitative and quantitative techniques of analysis. Dreher (2000) reviews previous efforts to use conceptual research in psychoanalysis (Sandler's Hampstead Index Project and Dreher's Trauma Project). Although this methodology is far from established within the field of psychotherapy research, analysts such as Tuckett and Dreher forecast that it will be an important part of hypothesis generation and testing for complex subjective concepts in the years to come (Dreher, 2000; Tuckett, 1994).

Future efforts using a similar study design

There are a number of ways in which a study with the quasi-experimental design chosen in this research could be conducted more effectively. First, more attention could have been paid to collecting a complete data set. This would have enabled useful statistical analyses of follow-along and follow-up data in a somewhat larger sample size. Second, a broader range of assessment measures could have been used, not limited to purely symptomatic measures (BDI, STAI-T) or counts of DSM diagnoses. In particular, the only convincing research to date that has shown a relationship between initial assessment and treatment outcome from psychoanalysis and psychodynamic psychotherapy comes from subdividing character pathology into introjective and anaclitic types (Blatt & Shahar, 2004). Therefore, it would have been useful to include some measures on which assignment to these two groups could have been made at intake.

3.5 Conclusion

In this chapter, the design and implementation of the Young Adult research scheme is presented along with justification for the use of quasi-experiments and a review of the relevant measures. The study was specifically intended to frequently collect

a broad range of outcome and process measures in a relatively small sample of young adults undergoing psychoanalysis and psychodynamic psychotherapy. The demographics, treatment parameters, and initial assessments were presented, revealing that the sample was relatively homogeneous and scored high on measures of depression, anxiety, and comorbid DSM Axis I and Axis II disorders. Associations between demographics, treatment, parameters, and initial assessments were few and in line with some predictions from the literature. Most importantly, a composite measure of improvement was described that uses the individual measures available in 19 of the 25 subjects at initial and termination. Sixty-three percent of subjects showed significant improvement, using this composite measure, 83% in the intensive group and 12% in the non-intensive group. There is some suggestion, in line with theoretical predictions and previous empirical research, that psychoanalysis is, in fact, a more successful treatment than psychodynamic psychotherapy, but this conclusion is limited by the small sample size, the longer average length of treatment in the intensive group, and the possibility of therapist allegiance effects. These results are discussed in the context of existing research and the advantages and disadvantages of different psychotherapy study designs.

The results from this chapter provide the beginnings of some answers to the questions raised in Chapter 1 and the introduction about the effectiveness of psychoanalysis. More impressively, though, they point to the limitations of studies that attempt to use contrast group methodology and the impossibility of collecting both a broad enough set of measures to characterize patients and how they change and also have a sample size large enough to make meaningful conclusions about the treatment in general. The solution to this dilemma that will be pursued in this thesis is to move from a focus on outcome only to a study of process and outcome together. This approach has been advocated widely in the psychotherapy research literature and has been used effectively to study the mechanisms of therapy in relation to change (Fonagy et al., 2001;

Greenberg, 1986; 1991). The following chapters will focus, in particular, on the YAWRS as a measure of process and the AAI as a measure of structural change, as they both relate to intensity of psychotherapy and treatment outcome.

CHAPTER 4. THE YOUNG ADULT WEEKLY RATING SCALE: MEASURE

DEVELOPMENT

The YAWRS is a versatile measure for gathering weekly data from the process of psychoanalysis and psychodynamic psychotherapy. First, the issues confronted in designing such a measure are discussed in the context of a review of existing process measures. Next, the development and revision of the measure are described as well as methodology for entering and correcting the data from this measure. Third, we describe several procedures, involving a combined theoretical and empirical approach, for reducing the many items of the YAWRS to usable and meaningful process variables and describe the results of these procedures. Finally, the results of these procedures are discussed in connection with prior process research and a new approach to answering process-outcome questions is presented.

4.1 Introduction

Ideally, the design of a psychoanalytic process measure is derived from a fit between the advantages and disadvantages of methodological strategies and the purpose for which the measure is being created. In proposing an individual measure, a researcher makes important choices regarding: (1) the underlying theoretical framework (e.g., general theory of common factors in psychotherapy vs. psychoanalysis), (2) the source of information (e.g., patient vs. therapist vs. external rater), (3) the nature and range of process information being collected (e.g., general information covering the full range of what happens in a session vs. specifics of a single process, such as transference interpretation), (4) the scope of material to which the measure is applied (e.g., individual words vs. several sessions), and (5) the nature of the material to which the measure is applied (e.g., subjective recall vs. videotape).

4.1.1 *Existing process measures*

A review of psychotherapy process methodology is remarkable for both the diversity of approaches and the large number of process measures, many of which collect heavily overlapping data. How a researcher quantifies psychotherapy process is influenced by perspectives on the important components of process and the most reliable and valid sources of information. The pendulum of theoretical and empirical answers to these questions has swung widely throughout the years. The field began with ostensibly atheoretical therapist and patient completed checklists, such as those developed by Orlinsky and Howard (1986), Eugster and Wampold (Wampold, 2001), Kolden (1996b), Weissman (1972), and Stiles (1980). While these measures were easy to implement, usually boasted good psychometric properties, and yielded a large literature of interesting findings, concern mounted that they were fundamentally biased by their reliance on participant information.

The measures that followed were predominantly observer-rated schemes for describing various elements of the psychotherapeutic process. Only two widely used and validated general measures of process emerged, the Vanderbilt Psychotherapy Process Scale (VPPS; Suh, 1986) and Psychotherapy Process Q-set (PQS; Jones, 2000), both of which use ratings of audio and videotaped sessions to describe fundamental properties of the patient-therapist encounter. As with the preceding patient and therapist questionnaires, these measures have attempted to be atheoretical, and in order to achieve reliability they emphasize concrete and observable utterances and behaviours. Observer-rated measures that have attempted to remain broad, yet become theoretically specific (Holland, Roberts, & Messer, 1998; Waldron, Scharf, & Firestein, 1997; Wilke, 1997) or collect data on complex psychological processes (Stiles et al., 1990), have been unable to achieve consistent reliability or validity.

The field's approach to these constraints on measure development has been to devise more methodologically narrow and theoretically well-defined measures of process. Psychoanalytic and cognitive science theories for the relation of a patient's language and affect to psychotherapeutic process is well-developed, and, in the hands of good researchers, has led to interesting and important measures. Ratings of language (Bucci, 1997b; Mergenthaler & Bucci, 1999; Spence, Mayes, & Dahl, 1994; Russell, 1986) focus on the most accessible element of the therapeutic process and help answer specific questions about how this changes during psychotherapy, though it is still unclear how they relate to underlying structures. Measures of affect (Bänninger-Huber & Widmer, 1997; Dreher, Mengele, & Krause, 2001; Hölzer, Pokorny, Kächele, & Luborsky, 1997; Horowitz, Ewert, & Milbrath, 1996; Krause, Steimer-Krause, Merten, & Ullrich, 1998) sacrifice the methodological advantages of language measures, for a focus on states of mind that feel more central to the human experience of therapy. However, successful research efforts with affective measures have been limited.

General investigation of therapist interventions, using both molecular and molar observer-rated measures, has been a major movement in the field of psychotherapy research, though it has yielded few reliable relationships between intervention strategies and outcome (Hill et al., 1988; Piper, Debbane, de Carufel, & Bienvenu, 1987). When theoretically specific interventions are targeted, the results appear to be more encouraging (Gaston & Ring, 1992; Milbrath et al., 1999). The most psychoanalytic-specific exploration of therapeutic process, focusing on observer ratings of transference interpretations, has been methodologically difficult but, in small samples, is already showing promise of elucidating mechanisms of change (Connolly, 1998; Bøgwald, 1999; Gabbard et al., 1994; Høglend, 1993c; Piper et al., 1991).

Interestingly, the most productive recent area of psychotherapy process research, has focused on the relatively atheoretical but highly reliable concept of therapeutic

alliance, as it is rated by patient, therapist, and observer. Studies have consistently shown that no matter how this variable is measured, but particularly when it is rated by patients, alliance is a strong predictor of positive treatment outcome (Bachelor, 1991; Marziali, Munroe-Blum, & McCleary, 1999; Piper et al., 1991; Safran & Wallner, 1991; Tichenor & Hill, 1989).

Despite all the work that has been done in this area, the question of how therapist, patient, and observer ratings of therapeutic process compare in their usefulness and predictive validity remains an unanswered one. All three perspectives have clearly been useful, under a wide variety of conditions and from a range of theoretical perspectives, in learning about the therapeutic process. Given that no evidence has emerged for one of these perspectives to be considerably superior, it is reasonable to conclude that all are potentially valuable and constraints of methodology and experimental practicalities will continue to dictate whether one or all three of these can be used.

On this basis we believe that a successful measure of psychotherapeutic process would be one that is theoretically well-defined, yet incorporates multiple theoretical and clinical perspectives, and collects ratings as directly as possible from the raw material of a psychotherapy session, yet is not overwhelmed by large quantities of difficult to code data. Language and affect measures do not satisfy these criteria because they are too clinically and theoretically narrow and, in order to cope with the vast amounts of data they yield, sacrifice large scale interpretations for an emphasis on microanalysis.

Theoretically focused observer-rated measures of therapist interventions and transference are promising but miss out on the more subjective elements of therapeutic experience and look at only one aspect of process. The PQS is probably the best existing observer-rated measure of general process, but in its quest for reliability and wide-applicability lacks items describing psychoanalytically specific phenomena.

This argument leads to the idea that perhaps, as Orlinsky and Howard originally suggested (1986), the therapist and patient are the best sources for accurate and meaningful information about the therapeutic process. Patient report, while clearly interesting, is fundamentally limited by the inability to describe psychoanalytically sophisticated and unconscious phenomena. Therapist reports, on the other hand, while reasonably the object of suspicion in any outcome study, are surprisingly well correlated with patient measures and may not be as biased as initially feared (Weiss, Rabinowitz, & Spiro, 1996). Furthermore, therapist ratings have the undeniable advantage of utilizing a psychoanalytically-informed rater who can integrate large amounts of data from a uniquely important perspective. The Young Adult Weekly Rating Scale, described in this chapter, is a therapist rating scale designed as a general measure of process which incorporates multiple theoretical and clinical perspectives, with emphasis on the psychoanalytic, and addresses the theoretical and methodological challenges described above.

4.1.2 Proposed structure of the YAWRS

In order to answer the broad range of questions about psychoanalysis raised in Chapter 1 and reviewed in Chapter 3, the YAWRS was designed to collect information about the entire gamut of psychoanalytic process including (a) the behaviour of a patient in a session, (b) the manifest content of the patient's report and how it is interpreted by the analyst, (c) the affect, transference themes, and defences that are contained in the patient's material and their interpretation by the analyst, (d) unconscious themes and their interpretation, and (e) the analyst's style of intervention and countertransference experience of the patient and session. Such a range necessitates a large and time-consuming process measure. In order to provide a manageable unit of analysis and prevent the analyst from being overwhelmed by data collection responsibilities, the YAWRS was designed to be applied to four or five sessions at a time (over the course of

one week in the case of psychoanalysis or a month in psychodynamic psychotherapy). Since audiotapes, videotapes, and transcripts were not available, the YAWRS was rated on the basis of subjective recall and an analyst's review of their own process notes.

The enormity of the psychoanalytic literature bears witness to the fact that there are many ways in which analysts organize concepts of process and link them to outcome. In the process literature, this problem has typically been handled by beginning with a large number of items describing specific elements of process and using some combination of theory and factor analysis to organize the items into useful scales. By showing which items tend to be scored in the same direction, factor analysis reveals how the raters organize their thinking. The resultant scales capture the broader concepts and minimise uncertainty by averaging multiple items whose errors are not correlated and therefore cancel each other out (Anastasi & Urbina, 1997; Nunnally & Bernstein, 1994). At the same time, theoretical beliefs about how items fit together are necessary to guide the development of scales that may not immediately hang together from observation alone. Each of the findings from the process-outcome literature suggests a link between an outcome finding and a *theoretical* construct of process which is measured differently depending on the process measure.

We propose to use both the empirical means of factor analysis and theoretical guidance from the psychoanalytic process and outcome literature to shape development of the YAWRS. A large number of items covering all the major areas of theoretical interest were compiled by expert raters with the plan of collecting a large data set and then applying factor analysis to explore the measure's structure. This chapter describes this method in detail and the resulting structure. In Chapters 5 and 6 we will compile interesting factors and items into theoretically meaningful scales for the testing of discrete hypotheses.

4.2 Methods

4.2.1 Subjects

The YAWRS was completed at various time points for 19 subjects (12 in psychoanalysis and 7 in psychodynamic psychotherapy) enrolled in the Young Adult Research Scheme. For 10 of these subjects (4 in psychoanalysis and 6 in psychodynamic psychotherapy) the YAWRS was collected consistently for more than half of the first year of treatment. Data analysis was performed on these 10 subjects. The demographic characteristics, initial diagnoses, treatment parameters, and symptom profile of the YAWRS subjects during treatment and follow-up are described in Chapter 3. Of the 10 subjects for whom first year YAWRS data were consistently available, 4 (subjects M, N, O, and U) were judged to be symptomatic improvers and 6 (subjects B, F, K, L, R, and Y) non-improvers by the criteria established in Chapter 5. Of the improvers, all but one (subject N) were treated with psychoanalysis, while of the non-improvers all but one (subject B) were treated with psychodynamic psychotherapy. On the basis of analyses reported in Chapter 3, it was not possible to say whether this association occurred by chance or due to an association between psychoanalysis and improvement in the larger sample. Other than this, no significant differences were found between the subsample used for the YAWRS analysis and the larger Young Adult sample.

4.2.2 Assessments

The Young Adult Weekly Rating Scale (YAWRS)

The Young Adult Research Group designed the Weekly Rating Scale (YAWRS) as a checklist for rating the presence of all possible themes that could emerge in a week of analytic sessions (see Appendix 4.1). Unlike existing psychotherapy process checklists the YAWRS was designed explicitly to include a wide range of psychoanalytic themes, but the authors were careful to provide items that would fit all orientations of psychoanalytic

theory and clinical practice (with particular attention to differences between the Contemporary Freudian, Kleinian, and Independent groups of the British Psychoanalytical Society). It was hoped that requiring analysts to rate each carefully described item would force them to identify the presence or absence of themes, even when they had not been explicitly aware of the status of these items beforehand. In the widely used free-form weekly, therapists have a chance to record the themes of the analysis on which they have been consciously focusing. The Young Adult Group believed, however, that the open format would allow them to leave out less conscious but perhaps equally (if not even more) crucial themes in the analysis. The YAWRS was designed to capture all themes with equanimity and thus decrease the under-representation of less conscious themes in depiction of the analytic process.

Between 1992 and 1994, two major revisions to the YAWRS were produced and several minor corrections applied. In the original version (printed versions A, A2, and B), the checklist had a total of 537 items grouped in nine sections: (1) general stance to the analysis, (2) manifest ideational content (body, self, past and future, parents and siblings, current life events and difficulties, sexuality, dreams, analysis, relation with analyst, and relationship themes), (3) affective content, (4) behavioural content (attendance, free association and behaviour in the analysis, co-operativeness, and sexual and aggressive behaviour), (5) manifest mental functioning, (6) analytic understanding (defences, conflicts related to the patient's body and drives, conflicts related to the self-object relationship and self-evaluation, conflicts surrounding object relations and sexuality), (7) analyst's non-interpretive stance (supportive, facilitative, clarifications, suggestive/counseling, directive, protective of analytic process, other), (8) interpretive interventions (repudiated affect, current state, past, self-evaluation, transference, style of interpretation, resistance, response to interpretation, and subjective experience of the analyst), and (9) analyst's judgement of the quality of the week. All items in sections 1

through 9 were rated as absent (0), present during the week (1), or a central feature of the week (2). Items in section 9, evaluating the quality of individual components of the analysis, were rated on a 5-point Likert scale consisting of very poor (1), poor (2), average (3), good (4), and very good (5).

In November 1992, in response to comments from analysts who had begun to use the YAWRS with their young adult patients, two major modifications to the YAWRS were carried out, producing the second major YAWRS version (printed versions C2, C3, and C) . First, it was found that evaluating 537 items, even only for a moment each, was too time-consuming a task for the therapist to perform each week. This was addressed by decreasing the total number of items by approximately 50, and, more importantly, by rearranging the items into a hierarchical structure such that most specific items on the questionnaire need not be filled out when the general theme under which they appear is rated as “absent.” The minimum core of items to be completed was thus reduced to 80, with each answer of “present” leading to 5 to 10 more items to complete. Second, it was decided that the original format of the questionnaire did not adequately discern whether the analyst had followed his/her observation of themes in the analysis with interpretations guided by those themes. Therefore, in the revision each specific theme was to be rated on both its presence in the week and its interpretation by the analyst.

The revised structure of the checklist now had three main sections, with multiple subsections: (1) general characteristics (resistance, general attitude to analysis, time keeping, patient’s behaviour in session, quality of analytic material, patient’s aggression and sexuality in the analysis, and maturity of patient’s mental functioning), (2) manifest content (body, self-esteem, historical material, relationship themes within family, current life events, sexuality, discussion of treatment parameters, general transference themes, predominant affective themes, predominant defences, predominant UCS dynamic themes, predominant reaction of patient to interpretation, and analyst’s style of

intervention), and (3) analyst's judgement of the quality of the week. Items at the head of each subsection were rated as "yes" or "no" to determine whether the subsection needed to be completed. Individual items were rated for presence in the week as absent (0), possibly present (1), or prominent (2), and for interpretation by the analyst as not taken up (0), taken up indirectly (1), or taken up directly (2). Items in the final subsection of section 2, analyst's style of intervention, were rated on only a single five-point Likert scale from absent (1) to extremely important (5), with 0 signifying not applicable. Items in section 3 were rated from very poor to very good, as in the previous version.

In October 1993, after further feedback from analysts using the YAWRS, it was decided that another major revision was required to add greater breadth of items throughout the questionnaire, with particular focus on manifest and unconscious themes. These changes (leading to printed versions D, E, and G) included the following: (1) A subsection on sessions missed by the patient was added to the general characteristics section. (2) The manifest content section from the previous version was subdivided into sections entitled manifest content, preconscious themes, and predominant UCS dynamic themes. (3) Subsections on relationship themes with friends, sexual relations, money and work, gender, age, and race issues, and adult identity were added to the manifest content section. (4) A subsection on changes of transference across the week was added to the preconscious themes section. (5) Subsections on unconscious content relating to the body, reaction of the patient to interpretation as feeling helped, negative reaction of the patient to interpretation, and main impact of the patient were added to the predominant UCS dynamic themes section. (6) Items were added to subsections throughout the questionnaire to include themes pointed out by analysts in the Young Adult Group as having been left out in previous versions. In total, approximately 350 items were added to the questionnaire. However, due to the hierarchical structure, the minimum core of required items increased only to 103, leading to a maximum of 899 completed items.

Finally, as part of this revision, the scales used to evaluate presence and interpretation of items were expanded to four points: for presence, an item was rated as absent (0), possibly present (1), present (2), or prominent (3), and for interpretation an item was rated as not taken up (0), taken into account in other interpretations (1), taken up directly (2), or a central interpretation of the week (3).

Analysts in the Young Adult Research Group were instructed to complete a YAWRS form for each intensive patient at the end of every week in which the patient attended at least one session. For non-intensive patients, analysts were asked to complete a YAWRS at the end of every month in which the patient attended at least one session. Group members reported that depending on the diversity of material in the sessions from the preceding week or month, completion of the YAWRS required between 30 and 60 minutes. Analysts also reported that they became faster and more efficient at filling out the YAWRS as they became familiar with the structure of the measure and knew where to find particular items.

Between January 1992 and November 1992, 182 YAWRS forms (versions A, A2, and B) were completed for the first 10 intensive subjects in the study. Because of the large differences between this and later forms, these data were not used for further analyses and not included in the main data set. From December 1992 through October 1993, 356 further YAWRS forms (versions C, C2, and C3) were completed on nine of these original subjects (one terminated therapy) and two additional intensive subjects. Between November 1993 and June 1997, 980 more YAWRS forms were completed on these 11 intensive subjects, as well as on an additional 11 subjects (two intensive and nine non-intensive). Data from one intensive subject was not used because of concerns over confidentiality. In order to have enough data to recognize meaningful trends, a minimum of 6 months of completed YAWRSs for each subject was required for use of that subject's YAWRS data. This eliminated YAWRS results from two non-intensive subjects

(subjects G and H) and one intensive subject (subject I). The remaining data set consists of 1314 completed YAWRS forms from 12 intensive and 7 non-intensive therapies.

Other assessments

Demographic characteristics, initial diagnoses, treatment parameters, and symptom profile during treatment and followup were collected according to the experimental protocol described in Chapter 3. Other measures were not used in the statistical analyses of the YAWRS structure and YAWRS data were not subdivided at this time by treatment intensity or treatment outcome. These other assessments will be addressed in combination with the YAWRS in later chapters.

4.2.3 Data processing

Data entry

Data entry for the YAWRS was performed with a digitizing tablet and custom designed software (Tablet Oriented Data Entry System, TODES) in order to increase the speed of data entry and minimise errors. Using this system, a completed form is first placed on the digitizing tablet and “anchored,” by pressing with a digitizing pen at three corners of the form, to register its position on the tablet with the software. The pen is then pressed on every item circled on the sheet, causing a corresponding entry to be made into the computer database. Each page is anchored separately and in order, matching the software’s version of the same form. Using this system, data entry time was reduced from 45 minutes per form, using standard keyboard entry, to 10 to 15 minutes per form. Following entry of all 1314 completed YAWRS forms, a printout was made of all entered data and compared against the original forms by a different group of data entry personnel. Corrections were then entered into the computer via keyboard.

Conversion of data from previous versions of the YAWRS

In order to combine data from printed forms C, C2, and C3 with forms D, E, and G, two accommodations were necessary. First, the scales from the earlier version were

readjusted to match the later scales: scores 0, 1, and 2 on presence and interpreted scales were converted to 0, 1.5, and 3. Second, variables from C, C2, and C3 were relabeled to match variable names in the later version. Items which were not found in the earlier version were labeled as missing data. All further analyses used the converted form of C, C2, and C3 data.

Data repair

Upon data entry it was discovered that a few errors were made and shortcuts taken by some of the analysts filling out YAWRS forms. First, analysts often filled out the items of a subsection without explicitly filling out the general presence versus absence question at the beginning of that subsection used as the gateway (part of the minimum core of required items). Second, some analysts filled in only the items in a subsection that were rated as one or higher, leaving blank items they intended to score as 0. In both cases, the intentions of the analysts were verified via personal communication and corrections made to the entered data. Finally, in the last three subsections of predominant UCS themes (analyst's style of intervention, predominant aspect of analyst's feelings, and main impact of patient), analysts were confused by the use of 0 to signify "not applicable" and 1 as "absent," and often used 0 when they intended to respond "absent." Since the rating "not applicable" was not needed for the data analysis, all 0's were replaced with 1's in these subsections.

Labeling of YAWRS items and factors

In order to manipulate the large number of YAWRS variables, a naming scheme was designed to label the original items of the questionnaire, as well as factors and summary scores that were later calculated from the items. The general syntax for the name of a YAWRS item is:

UVWXYz where U = overall section (G=general characteristics, M=manifest content, U=predominant UCS dynamic themes, J=quality of week)

- V = first order subsection (A, B, C, D, etc.)
- WW = where needed: second order subsection (01, 02, 03, etc.)
- Y = where needed: third order subsection (A, B, C, D, etc.)
- z = where needed: p=presence, i=interpreted

example: MF04Gp = manifest content, sexual relations, being separate and independent – presence (see Appendix 4.1, page 7)

Summary factors derived from YAWRS items are named by removing the code denoting the subsection level over which the data has been summed or averaged. If more than one factor has been derived from a given subsection, they are labeled “f1”, “f2”, “f3”, etc. The letter “o” (for “overall”) is added to refer specifically to subsection summary scores, that were not derived from factor analysis. Finally, the letter “p” or “i” is added if the sum or average refers only to “presence” or “interpreted” items. For example, PA04f2i refers to the second factor calculated by averaging “interpreted” items from the manifest content, sexual relations, relationship themes subsection.

When discussing the YAWRS data, standard terms will be used to refer to the different components of the analysis. They are:

YAWRS item = individual question on YAWRS form (scored as 0 to 3, 0 to 1, or 1 to 5)

YAWRS factor (FAC) = score calculated from YAWRS items on the basis of factor analyses; definitions and formulae in Appendix 4.3a

YAWRS subsection summary scale (SSC) = score calculated from YAWRS items to summarize a single subsection, not using factor analyses; definitions and formulae in Appendix 4.3b

4.2.4 *Factor analysis and factor calculation*

Factor analysis of all YAWRS items

Factor analyses with Varimax rotation were performed on various sections and subsections of the YAWRS using the S-Plus data analysis software version 3.3 (Statistical Sciences, 1995). First, an individual factor analysis was performed on each of the 45

YAWRS subsections containing four or more items. The first 41 of these analyses were performed on “presence” items only. The last 4 were performed on the single five-point scale used for the following subsections: style of intervention, analyst’s feelings, impact of patient, and quality of week. Factor analyses were restricted to YAWRS data from subsections in which at least one of the items had been completed (indicated by proper selection of the “gateway” question), and missing data within such sections was converted to 0/“no” (except for the sections using five-point scales, in which 3/“average” was used for missing data in the “quality of week” section, and 1/“absent” was used for missing data in the other sections). Factor analyses were also restricted to data from the most recent major version (printed forms D, E, F) as some of the individual items were not found on earlier versions. The number of factors requested in each analysis was determined by visual inspection of a scree plot (factors chosen up until the point of change in the slope of the eigenvalues plotted against factor number) (Tabachnik & Fidell, 1989).

Calculation of YAWRS factors (FAC)

Results of the factor analyses were used to design formulae to calculate 227 factors that summarize the dimensions captured in each subsection of the YAWRS. Items with loadings of +0.25 or greater were considered to be positive contributors to the factor, and items with loadings of -0.25 or lower were considered to be negative contributors to the factor. Factor structure was also adjusted to conform to theoretically meaningful interpretations of factors. The factors were then calculated from the mean of positive and negative contributors. Where factors consisted of items with identical scales each factor was given the same weight: +1 for positive contributors, and -1 for negative contributors. In factors with a mixture of presence items (0 to 3) and binary items (1/“yes” vs. “0”/no), binary items were multiplied by +3 or -3 before summing, and presence items were weighted +1 or -1.

Organization of YAWRS items into factors based on “presence” items was also used to calculate factors from “interpreted” items. This was done as an attempt to quantify the level of interpretation corresponding to the same theme in the analysis represented by the “present” factor. All interpreted items were weighted +1 regardless of the weighting used for the “present” factor.

For some subsections, factor analysis revealed little correlation between items, because items represented different, and non-overlapping, manifestations of the concept summarized by that subsection. In these subsections, a one factor model was calculated and the factor formula was calculated as a mean of all the items in that section. An item was given a weight of -1 if its meaning was opposite to that of the other items in that subsection, not on the basis of the loadings from the factor analysis (such loadings could be misleading in the case of non-overlapping factors, because one item may correlate with the absence of the other even if both contribute to the same theme). As part of this analysis, 14 factors were also calculated that used the information from factor analysis and correlation matrices to summarize more complicated YAWRS items. Factors GD_p, GD_i, GH1, MA1, MC02, MC03, MH02, MH03, MJ03, MK1, ML01_p, ML01_i, PB01, and PC01 were calculated according to the formulae listed in Appendix 4.3a.

Factors were first calculated from the most recent major version of the YAWRS (printed versions D, E, and G) which were used to conduct the factor analysis and contained all the YAWRS items represented in the factor formulae. Next, the same formulae were used to calculate these factors from the previous major version of the YAWRS (printed versions C, C2, and C3). Since one or more items from a factor were often missing on these earlier forms, means were calculated with a different denominator. For factors in which all the items were missing in previous versions of the YAWRS, the results were labeled as missing data for older forms.

The validity of the subsection factors was assessed by calculating alpha coefficients for each. An alpha coefficient of 0.5 or higher is considered to reflect a cohesive factor because it indicates that the factor contains at least 25% of the variance of the items used to calculate it.

Calculation of YAWRS subsection summary scales (SSC)

A separate set of 101 scales was calculated from YAWRS items in order to calculate single dimensions for most YAWRS subscales, largely irrespective of factor analytic results from the previous section. Unlike the factor scales, these summary scales were assigned a 0 when the subsection was left blank, and labeled as missing only when the subsection did not exist at all (as was true for some of the subsections in the early version of the YAWRS). For most subsections these summary scores consisted of means of all the items in the subsection. Separate summary scores were calculated for presence and interpreted items. The subsection factor analyses described in the previous section were used to design formulae for factors in three subsections only, felt to contain important distinctions among their items: PD (5 defence mechanism factors), UG (4 styles of analytic intervention factors), and UH (4 countertransference factors). In total, 44 of the subsection summary scales were drawn from “presence” items, 44 were drawn from “interpreted” items, and 13 were drawn from items not explicitly labeled as either. The systems described in the previous section for weighting items and combining early and late versions of the YAWRS were also used for calculation of the subsection summary scales.

Subsection summary scale factor analysis

In dealing with the large number of items on the YAWRS, a single round of factor analysis succeeded only partly in reducing the data to manageable form. Subsection factor analyses reduced the 899 items to 227 factors, while subsection summary scales reduced the 899 items to 101 scales. Given the hierarchical structure of the YAWRS, it was not

possible to factor analyse all 899 items at once because only a fraction of these items were ever completed at the same time. Replacing blank items with 0's or their equivalent would also not lead to a useful factor analysis because of the enormous skewing of the item distributions, and the tendency for a factor analysis to focus on combining items within the same subsection. A more manageable set of YAWRS factor could be derived, however, by performing a global factor analysis on some subset of the 227 factors or 101 subsection summary scales. In particular, a factor analysis of the "presence" subsection summary scales (or more exactly, non-"interpreted" summary scales, as scales that are neither explicitly "presence" or "interpreted" should be included) would yield information as to the overall structure in which analysts filled out the YAWRS, and may lead to useful therapy process variables.

Factor analysis of the 57 non-"interpreted" subsection summary scales on the most recent YAWRS version (n=958) yielded two factors. Factors were assembled by assigning a variable to the factor on which it had the loading with the greatest absolute value, as long as that loading was greater than 0.25 or less than -0.25.

The results of this factor analysis were used to generate formulae for two second order factors (SOF). A mean of assigned items (selected according to the procedure described above) was calculated for each factor; items with positive loadings were weighted +1 and items with negative loadings were weighted -1. Although the factor analysis itself was restricted to the most recent version of the YAWRS (because of the absence of several of the subsection summary scores in the earlier version), global factors were also calculated for YAWRS forms of the earlier version, altering the denominator of the mean calculation based on the number of available subsection summary scores.

4.3 Results

4.3.1 Distribution of YAWRS data

YAWRS was collected on 19 subjects (12 intensive, 7 non-intensive) according to the patterns shown in Tables 4.2a and 4.2b.

Subj	Quarter																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
A	0	0	0	0	0	0	0	0	3	9	8	8	13	8	9	6	12	11	0
B	13	11	10	7	15	0	0	0	0	0	0	0	0	0	0	0	0	0	0
C	0	0	0	0	8	7	12	8	5	10	9	9	8	9	4	9	8	7	0
D	0	1	9	12	8	13	9	11	10	7	12	10	6	12	10	9	8	6	4
E	0	0	6	11	12	7	9	9	12	7	10	11	9	1	2	0	3	2	1
M	1	11	10	10	10	10	8	18	0	0	0	0	0	0	0	0	0	0	0
O	13	9	6	8	7	11	3	8	11	7	7	10	9	8	3	6	0	0	0
P	0	0	0	0	1	11	10	8	12	11	11	6	13	11	12	18	0	0	0
Q	0	0	0	0	0	0	10	11	7	10	12	11	9	0	0	0	0	0	0
T	0	0	0	0	8	10	9	11	11	10	9	11	9	10	7	7	8	11	0
U	13	11	8	10	13	10	7	7	11	0	0	0	0	0	0	0	0	0	0
X	0	0	0	0	0	1	8	7	7	10	4	7	8	7	5	4	3	10	0

Table 4.1a. Numbers of YAWRS forms collected by quarter year from beginning of treatment in intensive subjects.

Subj	Quarter																		
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
F	2	1	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
J	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
K	2	2	3	3	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
L	3	2	2	2	3	0	0	0	0	0	0	0	0	0	0	0	0	0	0
N	0	2	2	2	3	1	2	2	3	2	2	4	0	0	0	0	0	0	0
R	2	3	3	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	0
Y	2	3	3	3	2	1	1	3	0	0	0	0	0	0	0	0	0	0	0

Table 4.1b. Numbers of YAWRS forms collected by quarter year from beginning of treatment in non-intensive subjects.

4.3.2 Results of factor analysis

Subsection factor analyses

Results of the subsection factor analyses are presented in Appendix 4.2a. Of the 45 factor analyses performed, 16 yielded a single factor, 9 yielded two factors, 10 yielded three factors, 9 yielded four factors, and 1 yielded five factors. A list of the 29 multiple factor solutions is presented in Table 4.2. Factor analyses generally validated the theoretical intent of the YAWRS items in that they aggregated along theoretically reasonable lines. For example, examples of resistance were divided into active (withholds information, lies, brings lists, stays with conscious meaning only, is intolerant or analyst's views, and doesn't listen to the analyst's interpretations) and passive (is silent, speaks vaguely, and is repetitive or boring) categories. Formulae for the calculation of the 227 subsection factors (for both presence and interpreted factors) are presented in Appendix 4.3a.

General characteristics

1. Examples of resistance: active, passive
2. Time keeping: difficulty leaving, other time problems
3. Maturity of mental functioning: regression, confusion/concrete, primitive boundaries

Manifest content

4. Body: pride, disgust
 5. Body topics: psychotic, damaged, narcissistic
 6. Self-esteem: positive, negative
 7. Affective tone of memories: anxious, angry
 8. Relationship themes with family: identification/independence, punishment, narcissism, Oedipal
 9. Relationship themes with friends: narcissism, belonging to group, threatened, separateness
 10. Relationship themes with sexual partners: fear of intimacy, fear of being unloved, fear of attack, wish for stability
 11. Affective reaction to event: full of affect, depressed/suicidal, not happy
 12. Gender, age, race: patient culture/race, analyst culture/race, analyst gender, analyst age
 13. Sexual material – content: perverse, sexual/homosexual fantasy, sexual life
 14. Sexual material – motivation: inhibition/anxiety, shame/ambivalence
 15. Affect discussing treatment parameters: Sad/angry, guilty/anxious
-

Preconscious content

16. Positive transference: love/erotic, idealization/identification, dependence
17. Transference w/anxiety: paranoid, fear of rejection, projected aggression
18. Transference w/competition and aggression: analyst is helpless, rivalry/victory
19. Transference w/resentment: derogatory to analyst, abandoned by analyst, loss of contact, dependence
20. Primitive emotional stance: fear/shame, manic grandiosity, existential anxiety
21. Predominant defences: projection/reaction formation/denial, isolation, splitting, regression/projective identification, externalisation

Unconscious content

22. General: narcissistic infantile objects, valued object, greed/envy
23. Reactions to aggression: destructive, anxious, masochistic, depressed
24. Sexuality: Oedipal, identity, masochistic
25. Self and self-esteem: low self-esteem, narcissistic, obliterated, self-criticism

Analyst reactions and behaviour

26. Analyst feels patient is helped: contained, confronted
27. Style of intervention: interprets object relations, interprets thinking, perception, behaviour, supportive, interprets defences
28. Analyst's feelings: attacked, inadequate/confused, loving/empathic, bored/cutoff
29. Impact of patient: bullying, narcissistic, analyst rejected

Table 4.2. Multiple factor solutions from subsection factor analyses.

Subsection summary factor analysis

Subsection summary factor analysis was based on subsection summary scores as defined in Appendix 4.3b. The results of this factor analysis are presented in Appendix 4.2b. The two factors correspond to clinically meaningful concepts of (1) resistance (explaining 12% of the variance and with a high coherence, $\alpha = 0.94$) and (2) clear unconscious themes and regression (explaining an additional 8% of the variance and with a moderate coherence, $\alpha = 0.67$). The subsection summary scores loading on each of these factors are listed in Table 4.3 and the formula for their calculation is presented in Appendix 4.3c.

Factor 1: Resistance (29 variables loaded)

1. General: resistance [GAop], negative attitude [GBop], bad behaviour in session [GEop], immaturity of mental functioning [GHop], lack of quality material [-GFop]
2. Manifest: lack of material about the body [-MAop], relationship with family [-MDop], relationship with friends [-MEop], or adult identity [-MJop]
3. Preconscious:

- a. Transference: with competition and aggression [PA4op], with resentment [PA5op], primitive [PA6op], lack of positive wishes toward analyst [-PA2op]
 - b. Emotional stance: lack of affect [-PC1op], lack of sadness [-PC2op]
 - c. Defences: regression/projective identification [PDf4p]
4. Unconscious:
- a. Reaction of patient: patient reports not feeling helped [UF2p, -UF1p], analyst feels patient reacts negatively [UF4p, -UF3p]
 - b. Style of intervention: lack of interpretation of object relations [-UGf1], lack of supportive interventions [-UGf3]
 - c. Countertransference: analyst feels attacked/disgusted [UHf1o], inadequate/confused [UHf2o], bored cutoff [UHf4o], not loving/empathic [-UHf3o]
 - d. Impact of patient: bullying [UIf1o], narcissistic [UIf2o], rejects analyst [UIf3o]

Factor 2: Clear Ucs themes & regression (15 variables loaded)

- 1. General: no variables loaded
- 2. Manifest: self-esteem [MBop], gender/age/race [MIop], sexuality [MKop]
- 3. Preconscious:
 - a. Transference: with anxiety [PA3op]
 - b. Emotional stance: primitive [PC3op], anger [PC4op]
 - c. Defences: lack of isolation [-PDf2p], lack of externalisation [PDf5p]
- 4. Unconscious: general [UAop], aggression [UBop], sexuality [UCop], self/self-esteem [UDop], body [UEop]
 - a. Reaction of patient: no variables loaded
 - b. Style of intervention: lacks interpretation of thinking/perception/behaviour [-UGf2], lacks interpretation of defences [-UGf4]
 - c. Countertransference: no variables loaded
 - d. Impact of patient: no variables loaded

Table 4.3. Global factors from subsection summary factor analysis.

4.4 Discussion

The results presented above broadly support the usefulness of the YAWRS as a therapist-reported measure of psychoanalytic process and suggest that it successfully draws from the strengths of existing measures, while introducing several new and important features. First, the YAWRS follows in the tradition of the most influential process measures by taking a general approach and aiming to collect data on every aspect of the therapeutic process so as not to restrict the results to those from a predetermined

theory of change. These areas include directly observable aspects of patient process (behaviour in the session, manifest content of verbal report, and affect), directly observable aspects of therapist process (interpretation of material and style of intervention), therapist's understanding of preconscious and unconscious themes in the patient's material (defences, transference themes, fantasies, view of therapy, alliance), and therapist's understanding of his/her own preconscious reaction to the patient (countertransference, view of therapy, alliance). The findings that (1) meaningful factors were derived from the subsection factor analyses throughout the questionnaire and (2) scales derived from the subsection summary factor analysis included components from every major section (see Table 4.3) support the hypothesis that each of these areas is relevant to a global understanding of change.

4.4.1 Importance of a general approach

These results are much in line with well-replicated findings using the two most widely used general purpose process measures, the VPPS and PQS. First developed in 1973, the VPPS was used as an observer method for studying the differences in psychotherapy process between groups of analytically oriented, experiential, and nonprofessional therapists and attempting to link these differences to psychotherapy outcome (Gomes-Schwartz & Schwartz, 1978; Strupp & Hadley, 1979). The VPPS underwent two early revisions in which both theoretical and empirical (derived from a factor analysis) constraints shaped the selection of subscales and items (Gomes-Schwartz, 1978; O'Malley, Suh, & Strupp, 1983; Rounsaville et al., 1987; Smith, Hilsenroth, Baity, & Knowles, 2003; Suh, Strupp, & O'Malley, 1986). Eight non-overlapping subscales are calculated from 80 Likert-type items, the first five of which are solely derived from patient items, and the last three derived from therapist items: (1) patient exploration, (2) patient participation, (3) patient psychic distress, (4) patient hostility, (5) patient dependency, (6) therapist exploration, (7) therapist warmth and friendliness, and (8)

negative therapist attitude. A range of studies using the VPPS have demonstrated the interrater reliability, internal consistency, and predictive power of the subscales (Suh et al., 1986; Windholz & Silberschatz, 1988). Inter-rater reliability was superior with audio and videotapes than with transcripts (particularly on therapist warmth and friendliness, negative therapist attitude, and patient hostility) suggesting that these scales make use of subtle and non-verbal cues.

The PQS (1998; 1999; Jones, 1997; 2000; 1993; 1991; 1998; 1993) also began with a large number of concrete and atheoretical statements about therapy process, and through factor analysis of observer ratings using audio and videotapes derived a set of 100 useful and reliable items. However unlike the authors of the VPPS, Jones began with a more explicit theoretical framework for the active ingredients in therapy (outlined in Chapter 2) and addressed the methodological weakness of Likert scales by having raters follow Q-sort methodology, whereby the numbers of items endorsed or rejected for a therapy segment are dictated by a normal distribution curve (Block, 1961/1978). Similar to the YAWRS, the flexibility of the PQS stems from the fact that items have been used individually, in theoretically defined groupings (either defined by the experimenter or by a panel of experts rating “ideal” sessions), or factor analysed into empirically-derived subscales. The original such factor-analysis yielded four subscales: (1) therapist acceptance/neutrality, (2) therapist interactiveness, (3) psychodynamic technique, and (4) patient dysphoric affect (Jones, Cumming et al., 1993).

The two global factors derived from the YAWRS factor analysis appear to capture similar themes as some of the empirically derived factors of the VPPS and PQS. The resistance factor is highly related from the perspective of theory and face validity to several of the VPPS patient scales (patient exploration, patient participation, patient psychic distress, and patient hostility) and one of the PQS scales (patient dysphoric affect). Given the large body of data supporting the usefulness of the VPPS and PQS

scales, this supports the claim that the YAWRS regression scale will also provide meaningful data about the therapeutic process and be related to psychotherapeutic outcome. Meanwhile, the YAWRS factor capturing unconscious themes and regression does not have any face correspondence with the YAWRS and PQS scales. We believe that this represents the failure of the general measures to capture these more psychoanalytic and subtle variables. On the other side, the two global YAWRS factors certainly do not have the range of the eight VPPS and four PQS scales. This is left to the numerous subsection summary scores and factors, which will be explored in later chapters.

4.4.2 Exploring psychoanalytic concepts in depth

The most significant advantage of the YAWRS over the VPPS and PQS is the depth with which it seeks to investigate psychoanalytically relevant aspects of therapeutic process. Several general measures of process have attempted to do this, but with relatively little success in developing a practical approach with good psychometrics. Stiles and colleagues (Stiles et al., 1990; 1992; 1991) developed a method for testing their “assimilation mode” of psychotherapy which proposes that a systematic sequence of changes in the representations of a problematic experience (feeling, idea, memory, impulse, wish, or attitude) is central to the therapeutic process. Therapy transcripts were parsed into topics, then into insights and problematic experiences, whereupon raters scored these experiences on an 8-point scale (Assimilation of Problematic Experiences, APE) ranging from “warded off” (0) to “vague awareness” (2), then to “understanding/insight” (4) and ultimately “mastery” (7). Application of this measure to a number of therapies showed the gradual progression over time toward higher APE scores in successful treatments. Wilke used a similar approach, applying qualitative analytic techniques to the process notes of a long-term analytic psychotherapy (1997). In both these approaches, however, there was significant difficulty in reliably parsing and

rating transcripts. In addition, the measure, though general to all aspects of the process, was closely tied to Stiles's theory of therapeutic change.

In an approach closer to that of the YAWRS, Waldron and colleagues developed 25 "analytic process scales" (APS) and rated them on a set of sessions from the beginning, middle, and end of two psychoanalytic treatments (1997; 1999; 2004). Factor analysis led to the clustering of these scales into six categories: (1) patient quality, (2) analyst quality, (3) patient participation, (4) analyst participation, (5) patient involvement, and (6) analyst involvement. They describe how changes in these factors capture important themes within the two patients studied and propose future studies for relating process variables to outcome. Though promising, an absence of wider reliability and validity data, combined with a lack of funding has prevented this method from achieving its stated goals. Slightly more encouraging, Holland and colleagues have proposed the Rutgers Psychotherapy Process Scale (RPPS) as a reliable and valid measure for assessing psychodynamic psychotherapy sessions (Holland et al., 1998). Preliminary results suggest that their eight Likert-type items (significant material, development of insight, focus on emotion, direct reference to therapist/therapy, new behaviour in session, collaboration, clarity and vividness of communication, and focus on self) can be rated reliably from transcripts and are highly correlated with VPPS subscales. The partial successes of the APS and RPPS support the usefulness of a detailed general approach to measuring psychoanalytic process.

The other area in which process measures have been developed that look at psychoanalytic concepts in some depth is in therapeutic interventions. Gaston and Ring (1992) developed the Inventory of Therapeutic Strategies (ITS) in order to create a measure that is reliable, describes and compares the unique and common technical ingredients of various psychotherapies, and, most importantly, predicts outcome independently of measures of alliance (which had already been shown to be a robust

predictor as described below). The measure consists of 19 items subdivided into three categories: exploratory strategies (“therapist addresses the patient’s problematic defences [4 items], emotions [4 items], and cognitions [4 items]”), supportive strategies (strategies for solving interpersonal problems, alternative solutions to problematic solutions, reinforced patient’s change), and work-enhancing strategies (therapist sought patient’s participation in setting goals, encouraged patient to self-disclose or self-reflect, addresses patient’s problematic contribution, or explained the value of therapy). Raters score therapy sessions by first identifying every therapist statement with one of the 19 items, and then summarizing the entire session (making the ITS a molar method) by rating each item on a Likert scale of 1 (no emphasis) to 5 (major emphasis). Reported reliability has been good.

Cooper and Bond created the Psychodynamic Intervention Rating Scale (PIRS) by modifying the ITS into a molecular measure and simplifying some of the intervention categories (Milbrath et al., 1999). Each therapist utterance is classified by a rater as one of two interpretive interventions (defence or transference) or one of eight noninterpretive interventions (acknowledgements, clarification, questions, associations, reflections, work-enhancing strategies, support strategies, or contractual arrangements). This measure was applied to a single hour from 20 brief psychodynamic psychotherapies with bereaved patients, along with measures of patient process and 5 month patient symptom followup. Reliability of the PIRS was excellent.

The YAWRS is superior to Stiles’s and Wilke’s APES approach because of the greater theoretical breadth and the attention to more specific psychoanalytic concepts. It resembles Waldron’s APS and Holland’s RPPS, but has far more items, suggesting that it will be useful for asking more complicated questions about psychoanalytic process. Finally, the YAWRS is not limited to the emphasis on therapeutic interventions as is the

ITS and PIRS. We believe it benefits from the psychoanalytic depth that these measures employ, but will ultimately be more comprehensive and useful.

4.4.3 Coherent theoretical framework

An aspect of the YAWRS that is rare in process measures is the extent to which the organization and phrasing of the items are tied to a coherent body of psychological and psychotherapeutic theory. The work of the Young Adult Group is built from the training of its members as analysts in the British Psychoanalytic Society's Contemporary Freudian group. Items of the YAWRS, though written to capture multiple psychoanalytic perspectives, are all definable by a reasonably homogeneous theoretical framework, based on the work of Sigmund and Anna Freud, clarified by contemporary theorists such as Joseph Sandler. Few other process measures can claim this level of theoretical integrity. Perhaps the only system with a similar focus is the set of language measures developed by Bucci and colleagues as an outgrowth of her Multiple Code Theory (1993; 1997a; 1997b; 2000; 2001). In this theory, the cognitive "referential process" is said to connect three systems of information processing: the subsymbolic nonverbal (dominated by somatic and sensory systems), the symbolic nonverbal (imagery), and the symbolic verbal (language). Emotional schemas, which are seen as the central target of psychotherapy, are made up of all three systems and can only be successfully modified if all three systems are engaged. The referential process can be divided into phases which occur repeatedly within a session and across a treatment. Individual phases are associated with operational indicators in language and behaviour, particularly Bucci's Referential Activity (RA) measure, which reflects the linking of nonverbal experiences to language. A great deal of empirical work has been done in perfecting observer ratings of RA and the development of a computer assisted method for tracking RA (CRA) via the transcript of a session (Mergenthaler & Bucci, 1999). Though impressive in their theoretical rigor and empirical faithfulness to this theory, Bucci's measures are not as generalizable as the YAWRS and

are restricted to the “symbolic verbal” domain, whereas the YAWRS collects data from all three systems of information processing.

4.4.4 Number and organization of items

Two aspects of the YAWRS are particularly unique in the literature of psychotherapy process measures: the large number and hierarchical organization of the items. The presence of 899 items provides the opportunity for collecting specific data on an enormous range of specific phenomena, without the need for asking the reporter to summarize or generalise their answers, operations that might lead to inter-rater variation and error. It also allows considerable overlap in responses, which after aggregation of appropriate items, can reduce error and provide a reliable summary scale. The hierarchical organization helps the analyst completing the YAWRS organize the data from the session and with repeated application of the measure, remember and plan where the item is that best captures the session content. It also reduces the number of items answered on a given questionnaire and makes it possible for the questionnaire to be completed in a reasonable 30 to 45 minutes.

These two aspects of the YAWRS make it particularly distinguishable from two therapist report measures that it otherwise resembles. Graff and Luborsky (1977) designed a 23 item therapist process checklist divided into seven categories: (1) patient (reflective, receptive, anxious, depressed, hostile, and other), (2) resistance, (3) dreams, (4) interpretations, (5) therapist (active, empathic, warm, directive, and feeling reaction), (6) good hour, and (7) miscellaneous (technical problems, symptoms, and therapeutic change). Though there were some initially interesting findings using this measure (summarized in the Chapter 5), its brevity made it too limited for general application. Baer and colleagues (Baer, Dunbar, Hamilton, & Beutler, 1980) enumerated 74 items in their psychotherapeutic process inventory for therapists which were reduced, by factor analysis, to four scales: therapeutic participating, resistance, directive support, and

dysphoric concerns. The lack of replication with this measure is also likely due to its limited range. Scales by Waldron (1999) and Holland (1998) have fallen victim to the same failing. No other psychotherapy process scale to date has used a hierarchical structure.

4.4.5 Importance of empirical clustering

The use of factor analysis to empirically cluster items of a process measure has been a central part of the most successful techniques. Most notably, the VPPS (Windholz & Silberschatz, 1988) and PQS (Jones, 2000) use factor analysis of items to develop clinically meaningful scales and to provide validity for the items collected. Following in this tradition, factor analysis of the YAWRS items yielded clinically meaningful subsection subscales and global factors, suggesting that the items were capturing relevant data about the application of psychotherapy and psychoanalysis. The large α scores and percent of variance accounted for by each of the factors further supports the validity of the measure.

4.4.6 Importance of theoretical clustering

Empirical clustering of psychotherapy process items alone has several important weaknesses. First, items that cluster together are by definition ones that the rater sees as frequently co-occurring in the same session. If multiple items represent alternative theoretical or dynamic expressions of the same concept, but do not necessarily co-occur, they will not be grouped together using this method. Second, as is demonstrated throughout the literature on self-report measures (Stone et al., 2000) statistical clustering favors concrete and easily operationalised questionnaire items which have the greatest test-retest reliability. A purely empirical strategy for clustering items would lead (and has led with many measures) to scales that are valid but capture superficial and the most easily observable aspects of the therapeutic situation. Therefore, as has been done in the case of the VPPS and PQS, factor analytic methods should be supplemented with

theoretical constructions. The methods developed for the YAWRS utilize the enormous number of items by empirically clustering them within subsections and on the global scale, as well as using theoretical principles to develop meaningful and subtle variables.

4.4.7 Operationalised and practical

The theoretical sophistication and detail in process measures must be tempered by the practical constraint that items are definable, various raters can agree on their definitions in multiple settings, and the technique for rating them is not too complicated to be applied in a large scale study. The YAWRS, despite its length and theoretical complexity, consists of items that use standard psychoanalytic terminology and should be understood by various analysts as representing the same construct. Although it has not been done with this measure, a manual for operationalising the individual items would not be difficult to construct and used by raters who want to verify that they are rating an item properly. In keeping items straightforward, the YAWRS follows in the footsteps of widespread measures such as the VPPS and PQS. The technique for completing the YAWRS is understandable and should not lead to error or a more lengthy process. In this regard, it is in sharp contrast with measures applied to transcripts such as the APE (Stiles et al., 1990) which requires a complex parsing of a text before ratings can be applied.

4.4.8 Standardisability and good psychometric properties

The YAWRS, like any practical measure of psychotherapeutic process, has a good potential for being developed into a standard method with good reliability and validity. Once a manual is designed, studies can be conducted to explore test-retest and inter-rater reliability of the measure. For example, analyst raters could be asked to rate the same video-taped session on multiple occasions, perhaps after being given other distracting clinical material about the same case in the intervening time, to demonstrate that the analyst is rating the session as viewed, as opposed to using other clinical material. Inter-

rater reliability would be measured by having a variety of analyst raters score the same videotaped session. Both methods present the problem that rating from a videotape is inherently different from rating a session in which one participated, particularly for a patient one has known for some time. However, the demonstration of good reliability would be an important step in motivating more widespread research with this measure. Tests of validity could easily be conducted by using the YAWRS to rate sessions simultaneously rated by established measures such as the PQS and VPPS. The ultimate test of validity is whether the YAWRS scales correlate with treatment parameters (e.g., frequency, duration) and measures of outcome. Preliminary results with the YAWRS that support its validity are presented in Chapters 5 and 6. In following this program, the YAWRS has the opportunity to follow in the steps of a number of successful process measures which have all demonstrated excellent standardisability and psychometric properties (VPPS, PWS, Orlinsky & Howard's TSR, and Eugster & Wampold's CSPSC).

4.4.9 Captures subjective experiences

In the context of recent psychotherapy process research, the fact that the YAWRS collects data only from the subjective report of the treating analyst is one of its most conspicuous advantages and disadvantages. Interestingly, the development and use of patient and therapist rated measures of psychotherapy process preceded most observer rated measures by several decades. Researchers have long suspected the accuracy of patient and therapist rating of psychotherapy because of the large potential for bias and the complicated interaction of perception and memory with transference, alliance, and other aspects of the therapeutic interaction. However, a number of researchers have also pointed out that by being involved in the relationship, therapist and patient have a context for their observations that are not reproducible with a third party rater. Furthermore, they have suggested that the reported bias is overestimated. Weiss and colleagues (Weiss et al., 1996) reviewed 41 studies of client-therapist ratings of process

and outcome, and found a great variability in the extent to which clients and therapists agreed. However, in all studies it appeared that both therapist and client reports correlated independently with objective measures of client improvement, supporting the idea that both perspectives offer useful information regarding the therapeutic process (Elliott & Anderson, 1994; Orlinsky & Howard, 1986). Some psychoanalytic researchers have gone so far as to say that the usefulness of therapist notes about a session exceed that of a recorded session, because of the synthesis and contextual reading that the therapist provides of the session (APsA, 1974; DeWald, 1972).

The first of the patient-therapist process measures, the Therapy Session Report (TSR), was developed by Orlinsky and Howard in the 1960's as a way to detect the effect of outside influences in the therapist's experience on session content (Orlinsky & Howard, 1967; 1986). Both the therapist (form T) and patient (form P) version of the questionnaire consisted of 145 items rated on a Likert-type scale organized broadly into 12 categories: (1) session topics, (2) client expectations, (3) client concerns, (4) client and therapist interaction, (5) therapist understanding, (6) therapist helpfulness, (7) client accomplishments, (8) client's motivation to return, (9) overall quality of the session, (10) client's level of functioning, (11) client feelings, and (12) therapist feelings. Therapists and clients completed the form immediately after a therapy session. Numerous studies have utilized one or both versions of the TSR and established its item reliability, factor structure, and clinical meaningfulness in tracking session process and client change (1996a; Kolden, 1996b; 1996; 1993; 1993; Orlinsky & Howard, 1986; Saunders, 1999).

Orlinsky and Howard argued that by collecting data directly from the therapist and patient, the TSR accesses information, unavailable to an external observer, about the subjective experience of the social relationship (Orlinsky & Howard, 1986). A factor analysis of 890 therapy sessions from 60 patients and 470 sessions from 17 therapists yielded seven dimensions of the intersubjective relationship and all but the first two were

derived from a combination of patient and therapist items: (1) patient agency vs. passivity, (2) therapist agency vs. catalysis, (3) “healing magic,” (4) ambivalent nurturance-dependence, (5) therapeutic alliance, (6) defensive impasse, and (7) conflictual erotization. The degree of covariance between therapist and client ratings on most of these factors suggests the sharing of the experience. Patient-therapist item correspondence was similarly good for concrete items related to dialogue, session development, and quality of communication. On the other hand, correlation between therapist and patient reports of the patient’s concerns, relatedness, and feelings was poor, indicating that both ratings are necessary to get the broader picture.

A number of other measures were developed in the 1970’s and 1980’s, like the TSR, to concretely quantify various elements of psychotherapeutic process through a therapist or client post-session checklist. Each was found reliable and used for one or two studies but was neither generalizable nor impressive enough to be used beyond its original research group. Weissman and colleagues (1972) developed a therapist questionnaire that asked for information on four dimensions: what the patient discusses, the quality of the patient’s talk, the overt affect expressed by the patient, and time. Eugster and Wampold (Eugster & Wampold, 1996; Wampold, 2001) designed an 80-item questionnaire, the Comprehensive Scale of Psychotherapy Session Constructs (CSPSC) for therapists and patients to evaluate nine components of psychotherapy process: (1) patient involvement, (2) patient comfort, (3) patient progress, (4) patient real relationship, (5) therapist involvement, (6) therapist comfort, (7) therapist expertness, (8) therapist interpersonal style, and (9) therapist real relationship. The quality of a session, as evaluated by both patient and therapist, was predicted by patient progress and involvement. In addition, therapist evaluation of quality was related to therapist expertness and patient evaluation was related to therapist real relationship (Eugster & Wampold, 1996).

Smith and colleagues (2003) adapted the VPPS into 42-item therapist and patient rating forms so as to have a short and reliable method for assessing their views of process. Separate factor analyses from a preliminary study of 40 patients in psychodynamic psychotherapy yielded six internally reliable scales for each measure: (1) therapist exploration, (2) negative relationship, (3) patient psychic distress, (4) patient participation, (5) therapist warmth and friendliness, and (6) patient dependency. Though promising, the discriminant validity of the scales (inter-correlations are high) and association with observer measures of process or outcome have not yet been demonstrated (Smith et al., 2003).

The YAWRS follows in the spirit of these measures in giving an exclusive focus to therapist ratings. There is ample evidence that this perspective yields valid, accurate, and clinically meaningful data about the therapeutic experience, and may not be as biased, even for outcome studies, as some have suspected.

4.4.10 Relation to language and affect

Another virtue of the YAWRS is its friendly relationship with two of the most theoretically interesting yet methodologically challenging recent areas of process research: language and affect. Language measures are said by their proponents to provide crucial and reliable data as to where in sessions particular aspects of process occur. Bucci's methods (described above in section 4.4.3) are among the most carefully developed and theoretically well-defined. Another language-related process measure has been proposed that attempts to measure the "analytic surface" through the co-occurrence of the pronouns "you" and "me" in the patient's speech (Mayes & Spence, 1994; Spence et al., 1994; Spence, 1998; Spence, Dahl, & Jones, 1993). Spence and colleagues claim that the more frequently such pronouns co-occur, the more the patient is jointly considering himself and the analyst in thoughts, fantasies, and plans. Mergenthaler and Kächele (1996) applied five computerized measures derived from the information theory and text

analysis literature (speech variability, redundancy, abstraction, part of speech distance, and emotion tone) to a pair of analytic cases and found related but distinct patterns in the change of these measures over time. They recommend that further work be done with single cases seeking to confirm the generalizability of these measures across psychopathologies, and that group studies be delayed until this is better understood. Canfield and colleagues (1991) applied a computerized rating of emotional, cognition, and contract to sessions of three different therapists with a single client and found that the measure recorded characteristic patterns of interaction unique to each dyad.

Measurement of patient and therapist affect is another approach to deriving process data directly from the raw data that is relevant to psychotherapeutic mechanisms. Krause's Saarbrücken research group (Dreher et al., 2001; IPA, 2001b) and Bänninger-Huber (1992; 1997) have applied Ekman and Friesen's Emotional Facial Action Coding System (EMFACS-7; Ekman & Rosenberg, 1997) to measure moment-by-moment affect in videotaped psychodynamic psychotherapy sessions. To optimally describe conscious and unconscious emotional states, the Saarbrücken group collects data both from the EMFACS, which associates specific facial motor patterns with the primary affects (anger, contempt, disgust, fear, happiness, sadness, and surprise), and a German version of the Differential Emotions Scale (DES) which asks therapist and patient to self-rate the same emotions after every session. Because both measures are related to concrete descriptions of affect, reliability (interrater on the former and test-retest on the latter) is excellent. Hölzer and colleagues (1997) combined the advantages of affective and linguistic measures, using a computerized method to measure the frequency of "emotion words" in session transcripts.

The greatest weakness of language and affect measures, as described above, is that they are applied to the microprocess of sessions. If they are scored by raters, this means that they are enormously time consuming and can only be performed practically on a

relatively small portion of the available data. If they are scored by computers, there is an inherent problem in capturing complex themes that require human or therapist judgement. The YAWRS includes items that refer directly to language and affect without suffering from either of these problems. Therapist raters provide a gestalt impression of the role of language and affect in the session using the sophistication of a clinician's approach, but on a large scale that is clinically meaningful and inclusive.

4.4.11 Relation to case-formulation and transference

In a similar fashion, the YAWRS provides direct data on case-formulation and transference without the intense labor-requirement, poor reliability, and narrow focus of these measures. Luborsky (1998) identified 17 measures (since 1976) in this category which share the goal of identifying recurrent and characteristic patterns in a patient's relationships, with the assumption that these are importantly linked to their current difficulties and the aim of a psychotherapeutic intervention. When viewed statically these measures provide the closest we have to a systematic psychodynamic formulation or diagnosis. When applied to the transcript of a psychotherapy session (which all of them are designed to do), they can be thought of as measures of transference relationships or psychotherapy process. When studied at multiple time points during a treatment, they can be considered to be psychodynamically-informed measures of structural change. Although the details of the individual measures and a full review of the large literature concerning their application to psychotherapy outcome is beyond the scope of this chapter, some identifying features of the seven most widely used measures and important studies that use them will be reviewed below.

The Core Conflictual Relationship Theme (CCRT), introduced by Luborsky in 1977, is one of the original and most widely used of the transference-related measures (Luborsky & Crits-Christoph, 1998). To score the CCRT a rater subdivides a transcript (usually from a psychotherapy session) into a series of "relationship episodes" and in

each of these identifies what the patient is wishing for (W), what the real or expected responses from others are (RO), and how the self responds to the other's behaviour in reality or fantasy (RS). Each identified W, RO, and RS is then assigned to the best-fitting of eight standardised categories (a different set of possible categories is available for W, RO, and RS). Data analysis of these results varies somewhat from study to study, but usually involves selecting the most prevalent W, RO, and RS in a transcript or calculating the pervasiveness (frequency of RE's containing a specific W, RO, or RS divided by the total number of RE's) of specific categories, chosen either theoretically or from past transcripts. Reliability for selection of standard categories has been found to be adequate (0.61 to 0.70), though it is difficult to find reported reliabilities for selection of RE's (IPA, 2001a).

Though the CCRT has been used widely as a measure of psychoanalytic process and to develop other measures (e.g., Central Relationship Questionnaire, a self-report measure of the CCRT), relatively little application has been made to the direct assessment of psychotherapy outcome (Barber, Foltz, & Weinryb, 1998; Luborsky, 2000; Weinryb, Barber, Foltz, Göransson, & Gustavsson, 2000). Crits-Christoph (1998) reports on a study of 33 patients in psychoanalysis or psychotherapy whose early and late sessions were subjected to the CCRT to detect change over time. The most frequent negatively valenced relationship theme (W, RO, and RS) from early sessions was selected, and found to be significantly less pervasive in late sessions. Meanwhile the most frequent positively valenced theme from late sessions was selected and found to be less pervasive in early sessions. The pervasiveness of individual positive or negative themes were found to be correlated with one another and with standardised measures of symptomatology (SCL-90 and HSRS) in the directions predicted. While these results are interesting and appear to support the theory that the CCRT captures the role of psychotherapy in improving transference patterns, they suffer from the methodological weakness that the

change in pervasiveness may have been due more to the random phenomenon of regression to the mean than to a treatment effect. Further studies using the CCRT as an outcome measure, particularly in single case studies, have been conducted and these methodological flaws are being addressed (IPA, 2001d).

Perry and colleagues built upon their work in the systematic description of psychodynamic defences (Perry, 1989a; Vaillant, 1971; 1993) to design a tool similar to the CCRT known as the Idiographic Conflict Formulation method (Perry, 1989b; 1994; 1989). Drawing from an unstructured interview, psychotherapy session, or similar material, a well-trained clinician identifies a set of conscious and unconscious (1) wishes, (2) fears, (3) symptomatic and avoidant outcomes, (4) vulnerabilities to specific stressors, and (5) best level of adaptation to conflict. The first three of these form the core of the dynamic formulation and the first two can be matched to a set of 40 standardised wishes or 39 standardised fears. Reliability of these ratings has been found to be adequate and one study has shown a correspondence between the wishes identified by the ICF and those by two other case formulation measures (CCRT and Plan Diagnosis, Perry, Luborsky, Silberschatz, & Popp, 1989). To date no significant work has been published which demonstrates the reliability or validity of the ICF as a measure of psychotherapeutic change.

The Plan Formulation Method (PF) was developed at Mount Zion Hospital as an operationalisation of Weiss and Sampson's Control-Mastery Theory, a cognitively oriented psychoanalytic approach positing unconscious pathogenic beliefs about relationships that are acquired in childhood and reproduced in the therapeutic relationship until they are disproven by the treater (2003; Weiss & Sampson, 1986). In this procedure, clinical judges trained in the Mount Zion theoretical orientation identify four component parts of a formulation: (1) conscious and unconscious goals for therapy, (2) pathogenic beliefs that interfere with reaching those goals (obstructions), (3) specific

plans employed by the patient in the treatment to disconfirm these beliefs (tests), and (4) insights presumed necessary for improvement (Curtis, Silberschatz, Sampson, & Weiss, 1994). In order to make these assessments, judges independently compose lists for each component, combine these into master lists, rate all the items on the master lists for their relevance to the case, and then reach a consensus on which highly related and inter-judge reliable items should be included in the formulation. Reliability, calculated in this way, has been excellent, and several studies have demonstrated that the degree of adherence of therapists to the individual's plan formulation predicts good patient outcome. Furthermore, a measure called "plan attainment" that rates the degree to which the patient has achieved the goals and insights and overcome the obstacles outlined in the plan formulation, is a valid and useful measure of psychotherapy outcome. Finally, considerable work has been done within the Mount Zion group to test individual hypotheses regarding treatment interventions using the plan formulation approach (1994; Weiss & Sampson, 1986). Thus, while limited to a relatively narrow theoretical framework, the PF method has been an excellent illustration of how a well-defined measure can be useful in evaluating clinical psychoanalytic theory.

Configurational Analysis (Horowitz, 1993; 1989; 1994; 1995; 1995) is an ambitious scheme for identifying (1) recurrent patterns of experience or behaviour termed "states of mind", (2) interpersonal role-relationship models (RRM), and (3) information processing patterns based on a set of clinical material, with a particular focus on psychiatric signs, symptoms, maladaptive traits, or interpersonal problems. RRMs identify schemas (structures of meaning that affect thinking, planning, and action concerning the self and other), given them an organizing framework, and present them diagrammatically. RRMs could be useful as research tools, helping conceptualize patients at the start of therapy and detecting whether they have changed. However, little empirical

work has been done on the reliability of these assessments and they have not yet lived up to their potential (Horowitz & Eells, 1993).

The Consensual Response Method (CRM) is the most loosely defined and theory-neutral of the case formulation methods (1994; Horowitz, Rosenberg, Ureño, Kalehzan, & O'Halloran, 1989; Rosenberg et al., 1994). A set of clinical judges watch a semistructured interview of a patient, generate dynamic diagnoses, discuss and rewrite their diagnostic profiles, and finally count up the “thought units” in their profiles and retain those “consensual responses” that occur most frequently. Reliability in the thought unit extraction procedure was found to be high (though none is reported on other elements of the process). Some validity has been demonstrated by successfully using the CRM to predict problems discussed in therapy (as measured by the Inventory of Interpersonal Problems, Horowitz et al., 1988) as well as therapy outcome (patients with higher interpersonal content in their formulations did better in therapy). However, the theoretical ambiguity of the measure and great reliance on well trained and consistent judges makes it unfeasible for many studies.

The SASB-CMP derives from two research traditions: Benjamin's standard model of interpersonal functioning (Structural Analysis of Social Behaviour, 1974; 1994; 1996) and the Vanderbilt psychotherapy research group's identification of “cyclical maladaptive patterns” (Henry, Schacht, & Strupp, 1986; Johnson, Popp, Schacht, Mellon, & Strupp, 1989; Schacht & Henry, 1994). Raters trained in the SASB and CMP draw from a clinical data source information in three domains: (1) interpersonal actions and reactions, (2) internalised responses directed toward the self (“introjective acts”), and (3) fantasies or expectations, composed of predictions, wishes, and fears that could address either interpersonal or introjective phenomena. This information is recorded using well-validated and standardised SASB codes and may be used to generate a “causal hypothesis” to organize the preceding information into a cyclical interpersonal pattern.

Unlike other methods, such as the CCRT, the SASB-CMP relies on raters judgement of important themes, as opposed to frequency of occurrence within the session. Since the coding is based on an already reliable coding scheme, the reliability of SASB-CMP is believed to be excellent with well trained coders. Validity has been demonstrated in good correspondence with the CCRT (Johnson et al., 1989) and association with DSM-III/IV personality disorder diagnoses (Benjamin, 1994). The measure's authors point out its theoretically high sensitivity to therapeutic change and the use the SASB-CMP as a measure of structural change is being explored.

FRAMES (Fundamental Repetitive And Maladaptive Emotion Structures) is a tool for assessing psychopathology, the therapeutic process, and treatment outcome using a basic classification system for emotions and a theory for how they function (Dahl, 1991; 1998; 1994). The emotion theory has four propositions: (1) emotions share the properties of somatic appetites, such as hunger and sex, (2) one class of emotions, termed "It" emotions have objects and function as appetitive wishes about those objects, (3) a second class of emotions, termed "Me" emotions, function as beliefs about the state of fulfilment and nonfulfilment of these wishes, and (4) It and Me emotions together form a feedback system that provides information about our fundamental motives and their outcomes – fulfilled wishes lead to Positive Me emotions and unfulfilled wishes to Negative Me emotions. FRAMES are identified in a therapeutic session by four possible methods: (a) a prototype is formed from one narrative and the rater looks for instantiations in other narratives, (b) a prototype is formed as a generalisation of several narratives and substantiations are enumerated, (c) randomly arranged "idea units" from a transcript are translated into "prepositional units" and then coded using a standardised list of actions, times, and objects, or (d) an "object map" is constructed from the patient's narrative and standardised emotions are assigned to each of the FRAMES identified. Multiple research groups have made use of the various methods for assigning

FRAMES and, with great effort, achieved good reliability. However, the research is still at an early stage and no direct application to structural change has yet been published.

The YAWRS includes elements of all seven of the measures presented above without the time-consuming methodology and difficulties with reliability that they entail. The measures are also quite specific to patient structures and neglect almost entirely the role of therapist intervention in the evolving process. As significant data, sometimes on the same cases, has been collected using the case-formulation and transference measures, it would be useful to apply the YAWRS to these same sessions and confirm the likely relationships between these approaches.

4.4.12 Captures therapist intervention and interpretation

The YAWRS has the greatest range of all the general process measures in capturing therapist interventions and interpretations as they are related to the content in a series of sessions. No other measure matches scores on content so directly with the extent of interpretation. In doing so, however, it borrows from several existing measures of therapist intervention. Gabbard and colleagues (Gabbard et al., 1994; Horwitz et al., 1996) set out in the Menninger Treatment Interventions Project (TRIP) to test theoretical predictions about the effectiveness of transference interpretations in the treatment of three patients with borderline personality disorder using a seven category molecular measure of therapist interventions. A team of three clinical judges identified each therapist statement as belonging to one category on a continuum from most expressive to least expressive: interpretation, confrontation, clarification, encouragement to elaborate, empathic validation, advice or praise, and affirmation. In addition, each statement was classified as being transference or extratransference in nature.

Transference interpretation was defined as “an explanatory statement focused on the patient’s feelings, attitudes, and behaviours toward the therapist and linking two or more elements into a new relationship” (Gabbard et al., 1994, p. 63). Further assessments were

made by separate raters on predictions from initial material regarding the extent to which supportive and expressive measures would be used, an assessment of the patient's collaboration with the therapist, and an overall assessment of the changes made by the patient by termination and followup.

Another approach to devising a reliable and meaningful transference measure has been to use molar measures based in concrete phenomena. Bøgwald and colleagues (Bøgwald, Høglend, & Sørbye, 1999) used clinicians with extensive training in psychodynamic therapy to score a 15-item scale, called the Specific Therapeutic Technique (STT) scale. The first five items are devoted to transference interventions and are rated on a 5-point Likert-scale ("not at all" to "very much"): (1) focus on patient/therapist relationship, (2) questions about patient's feelings toward therapist/therapy, (3) interpretation involving patient/therapist relationship and one or several dynamic elements (impulse, anxiety, defence), (4) questions about patient's thoughts about what therapist might feel toward patient, and (5) linking of repetitional patterns (including genetic interpretations). When this method was applied to two sessions each from 30 subjects randomly divided into two manualized forms of brief dynamic psychotherapy, scale inter-rater reliability was excellent. This study also showed partial overlap with a recent measure of therapist interventions, called the Interpretive and Supportive Technique (ISTS) scale (Bøgwald et al., 1999). This 14-item scale includes three items specific to transference interpretations: (1) direct attention to the patient's subjective impression of the therapist, (2) make links between the patient's relationship with the therapist and the patient's relationships with others, and (3) focus on the patient and therapist in the treatment situation rather than the patient and significant others outside the treatment situation (Ogrodniczuk & Piper, 1999). Preliminary data on this scale shows it to be reliable and related to other existing therapist intervention scales such as Piper's TIRS (1987) and simpler Perception of Technique Scale (unpublished).

Both these measures have promise and await further studies to see how well they relate to outcome.

There has also been an effort to design a method to differentiate the diversity of dynamic interventions, with particular attention to different types of interpretations. Piper and colleagues (Piper et al., 1987; Piper, Joyce, McCallum, & Azim, 1993) developed the Therapist Intervention Rating System (TIRS) with 3 major classes and 10 categories of interventions: (1) noninterventions. (A) Noninterpretive interventions that *do not* refer to part of the patient's experience: (2) formal interventions, (3) information providing, (4) information requesting, and (5) directive. (B) Noninterpretive interventions that *do* refer to part of the patient's experience: (6) nondynamic component interventions. (C) Interpretations: (7) single, (8) double, (9) triple, and (10) quadruple component interventions. Six basic components may be included in an interpretation and determine the classification category: four dynamic: impulses, anxiety, defences, and dynamic expressions (affects, behaviours, and cognitions presented as part of an internal conflict), and two nondynamic: objects (persons other than the patient) and resultant expressions (affects, behaviours, and cognitions presented as end states). While appealing in its careful elaboration of different levels of therapist interventions, the complexity of the system and the absence of a clear research goal has made the method too cumbersome to achieve wide use.

The YAWRS shares elements with all these measures of therapist intervention in a way that is both more general and makes the link between patient material and therapist intervention more concrete. Research comparing the YAWRS therapist intervention items and scales to those of other measures would be useful in validating this part of the measure.

4.4.13 *Captures alliance*

No contemporary measure of process is complete if it does not address the variable that has led to the most replicable and robust finding in the process-outcome literature, namely the link between therapeutic alliance and patient improvement. A full review of the therapeutic alliance literature, both in terms of methodology and process-outcome results, is beyond the scope of this chapter. However a limited review of the major alliance measures is useful in putting alliance into perspective relative to the YAWRS. Four alliance measures make up the bulk of research in this area: (1) the Vanderbilt Therapeutic Alliance Scale (VTAS, Gomes-Schwartz, 1978; Hartley & Strupp, 1983; Moras & Strupp, 1982), (2) the Working Alliance Inventory (WAI, Horvath & Greenberg, 1986; 1989; 1991), (3) the Penn Helping Alliance Questionnaire (HAQ, 1996; Luborsky, Crits-Christoph, Alexander, Margolis, & Cohen, 1983; Morgan, Luborsky, Crits-Christoph, Curtis, & Solomon, 1982), and (4) the California Psychotherapy Alliance Scale (CALPAS, Gaston, 1990; 1991; 1988; 1998).

The first designated scales of patient-therapist alliance were part of the VPPS questionnaire and were designated by Gomes-Schwartz (1978) as “patient involvement” and “therapist-offered relationship.” From this, Hartley and Strupp (1983) developed the VTAS, a 44-item observer rated questionnaire, organized into six empirically-derived factors: (1) positive climate, (2) therapist intrusiveness, (3) client resistance, (4) client anxiety, (5) client motivation, and (6) client responsibility. Outside of the Vanderbilt group or a few studies comparing various measures of therapeutic alliance, little work has been done with this measure.

The WAI was developed by Horvath and Greenberg (1986), based on the theoretical work of Bordin (1979), as a measure of three aspects of the therapeutic relationship: (1) an agreement on goals, (2) the degree of concordance regarding tasks, and (3) the development of personal bonds. The original 36-item version collected data

independently from client and therapist, while later versions were developed with fewer items (Tracey & Kokotovic, 1989) and with versions for an independent observer (Andrusyna, Tang, DeRubeis, & Luborsky, 2001; Tichenor & Hill, 1989).

Based on a similar theoretical perspective, Luborsky and colleagues developed a system of “helping alliance” measures including the global rating method, the counting sign method, and an 10-item patient self-report questionnaire (Luborsky et al., 1983; Morgan et al., 1982). The questionnaire was expanded to 19 items and applied to both patients and therapists, showing good internal consistency and test-retest reliability (Luborsky et al., 1996).

The CALPAS began in the 1980’s as the Therapeutic Alliance Rating System (TARS), a 41-item questionnaire for clinical judges consisting of four theoretical dimensions: patient’s positive and negative contributions and therapist’s positive and negative contributions (Gaston, 1990). Parallel versions forms of the TARS to obtain ratings from client and therapist perspective, as well, were developed by Marziali and yielded two broad alliance dimensions: positive and negative (1984). Marmar and colleagues (1989; 1989) renamed the measure the California Psychotherapy Alliance Rating System (CALTARS) and used factor analysis to derive five factors: (1) patient commitment reflecting the therapeutic alliance, (2) patient working capacity reflecting the working capacity, (3) therapist understanding and involvement, (4) patient hostile resistance, and (5) therapist negative contribution. Good internal consistency and correspondence with treatment outcome were found. Finally, in line with the influence of Bordin’s work on the WAI and HAQ, items reflecting agreement on goals and tasks were added to the CALTARS, yielding the 31-item patient and therapist completed CALPAS. Factor analysis of the new measure yielded a single general factor in the patient form, but five factors on the therapist form: (1) patient commitment, (2) patient working capacity,

(3) therapist understanding and involvement, (4) disagreement on goals and strategies, and (5) therapist negative contribution (Gaston, 1990).

The YAWRS includes alliance in multiple ways. An item at the end of the questionnaire, entitled “Patient’s general stance (therapeutic alliance)” captures an overall construct, while numerous items throughout the questionnaire capture elements of commitment, quality of material, relationship to the analyst, and resistance. In fact, both global factors produced by the factor analysis may be seen as having direct relationships to therapeutic alliance. The diversity of these approaches in the YAWRS will hopefully allow a full examination of the relationship between alliance and other process measures, as well as the treatment parameter and outcome data. Studies looking for direct correlation between YAWRS scales and well-established dedicated outcome measures, such as the VTAS, WAI, HAQ, and CALPAS, would be useful.

4.4.14 YAWRS weaknesses

The greatest weakness of the YAWRS stems directly from the exclusive focus on therapist rating of session patient content and therapist intervention. While this focus, as described above, made it possible to have expert rating on a wide range of clinical material, it eliminates several other large and useful sources of information. First, though it is likely that therapist report is not as biased as was once suspected (Weiss et al., 1996), there is no doubt that patients and third-party raters provide distinct information valuable to any process-outcome study. Ratings by a therapist are heavily influenced by transference and countertransference interactions in a way that even a well-intentioned therapist is not aware. In addition, the causal relationships between therapist description of patient content and outcome are inherently impossible to disentangle. One is always faced with the question, is the therapist describing something about the patient that is predictive of outcome, or is the fact of the therapist’s judgement itself causing the outcome. Patient and observer-based reports are necessary to shed more light on these

correlations. In using only the therapist as a source of process information, the YAWRS is unlike most mainstream measures of psychotherapeutic process, such as the VPPS and PQS.

The extraordinary comprehensiveness of the YAWRS is both one of its greatest advantages and a significant liability. The large number of items makes the questionnaire time-consuming and unpleasant for therapists to complete, and places a significant barrier in conducting studies with application of the measure over long periods of time and to large numbers of patients. The data in this study were collected by paying therapists to spend an extra session a week completing the questionnaire. Most studies do not have the resources to pay their therapists in this manner. Shortening the YAWRS on the basis of the analyses presented in this chapter is possible – specifically using the subsection factor analyses to create summary items that are scored once. Findings of later chapters could be used to identify the variables most closely associated with treatment parameters and outcome, and possibly suggest which items are less useful and could be eliminated. Both these plans have the problem of forming a new measure that then needs to be extensively tested, undoing some of the work of this study. In addition, it is difficult to predict whether shrinking the number of items will increase the error and change the overall usefulness of the questionnaire.

Despite the relevance of the YAWRS to language, affect, and case-formulation or transference themes, the measure relies on the therapist's global judgement of these concepts and provides no direct measure. Measures of language (Bucci, Mergenthaler, Spence, and Russell), of affect (Krause, Dreher, Banninger-Huber, Holzer and Horowitz), and of case formulation (Luborsky, Dahl, Perry, Horowitz, and Weiss) all contribute important information because of their direct measurement of in-session behaviour via videotapes and text. These measures also provide a molecular approach, to which the YAWRS does not have access. As described above, the precision of these

measures has great costs in terms of labor requirement and lack of global perspective, making them less suitable for the process-outcome tasks for which the YAWRS was designed.

Finally, as of yet the YAWRS has no reliability data, an essential requirement for any widespread measure of psychotherapy process. Though we have argued that the specificity of the individual items and the theoretical coherence of the psychoanalytic model suggest that reliability will be good, like-minded psychoanalysts are infamous for disagreeing on the meaning of even the most basic clinical material, and empirical verification of YAWRS reliability is sorely needed. Even before this happens, the individual items of the YAWRS should be operationalised, so that raters can formally learn how to understand what they are rating. There is a considerable literature of psychoanalytically-informed measures (e.g., Waldron, Wilke, Holland, and Stiles) that have been promising but never adequately demonstrated reliability. It will also be important to better understand how the number of sessions to which the YAWRS is applied and the length of time after the sessions that the questionnaire is completed affects the quality of response.

4.4.15 *Study weaknesses*

The most significant weakness of this preliminary study on the YAWRS is the small sample size and inconsistent times and durations of treatment for which the YAWRS was recorded. Although more than 1,300 rating scales were collected, the stage of treatment sampled and the number of questionnaires per subject varied widely. In order to maximise the data set, factor analyses combined different versions of the YAWRS and their results were likely skewed towards the therapist-patient pairs with larger numbers of questionnaires. Numbers of subjects were too small to perform confirmatory factor analyses within treatment, diagnostic, or improver divided subgroups. In a larger and or more consistent data set, it would have been helpful to

perform separate factor analyses on different therapist-patient pairs. The number of subjects and consistency of data collected also prevented other potentially useful data analyses. For example, no factor analyses were performed on the interpretation items. Time series analysis, a powerful technique for looking at temporal patterns within process variable which would be well suited to consistently collected YAWRS data, was not attempted. The results of any factor analysis are better when the ratio of questionnaires to items is large (Kazdin, 1992). Even 1,314 is relatively small when conducting a factor analysis of so many items and subsection scales.

4.4.16 Future research

As this study consists of an exploratory first pass at use of the YAWRS in a small process-outcome study, further research would be useful in almost every facet of the study to develop the methodology. A first step would be to shorten the YAWRS so as to make it more easily applied by clinicians to a larger number of analyses and psychotherapies. The chief difficulty in shortening the measure is that scales composed of larger number of items appear to be more useful in the analyses and have better psychometric properties. An attempt should be made to cut down on items in those areas that are already sufficiently well-represented (e.g., specific unconscious themes related to relationships and sexuality), and perhaps expand some of the less well represented areas including items related to alliance, contract discussion, and specific exploratory or supportive interventions.

Further work is needed to test the psychometric properties of the items and scales both within a single individual and across subjects. Individual subjects should be asked to complete rating scales on individual days of a month and then retrospectively after a week or month, so as to check test-retest reliability and the accuracy of retrospective accounts. Standardisation of the scales across different therapists and patients is difficult given the uniqueness of every patient-therapist interaction. An approximation could be

made, however, by asking a group of therapists to complete rating scales based on a set of audio or videotaped sessions completed by a standard therapist. Such standard sessions could also serve as the basis for test-retest reliability measurement.

Third, as has been demonstrated in the existing process-outcome literature, a study using the YAWRS would benefit immensely from simultaneous collection of data using existing patient and observer-rated measures. A set of audio or videotaped analytic and psychotherapy sessions could be subjected to one or more of the standard patient ratings of process as well as some of the observer-rated measures, including the VPPS, Jones PQS, and therapist intervention scales (e.g., TRIP, PIRS, IRS). In combination with one or more of the standard alliance measures collected via therapist, patient, and observer ratings, this data set would be key to judging the validity of the YAWRS and bringing together the process-outcome literature.

Finally, the nature of the study itself could be expanded by using a larger number of subjects, particularly with a more representative combination of psychoanalyses and psychotherapies each with positive and negative outcomes. In such a sample it would then be possible to factor analyse items and scales of the YAWRS, measure the internal relationships of the scales, trace the temporal development of the scales, and study relationships with other process measures all in groups subdivided by diagnosis, treatment parameters, and outcome. These analyses would undoubtedly shed light on the quality of information provided by a good measure of psychoanalytic process.

4.5 Conclusion

In this chapter, the YAWRS, a new comprehensive psychoanalytically-informed therapist-report process measure, was presented along with data supporting its theoretical relevance and clinical validity. Subsection factor analyses using a set of 1,314 completed rating scales revealed theoretically meaningful factors and suggested that this

measure is an effective way to collect psychoanalytic process data. Detailed comparisons were drawn between this measure and existing process measures, and an argument presented for the need for such a measure in our field. Despite several weaknesses in this measure and study (including complete reliance on therapist-reported process, small sample size, and the absence of psychometric data on the YAWRS), the results are an original and important step in the development of a useful process measure for psychoanalytic research. Much research is left to do in this area, and this study suggests a number of important directions to be pursued.

CHAPTER 5. THE YOUNG ADULT WEEKLY RATING SCALE: PSYCHOANALYSIS VERSUS PSYCHODYNAMIC PSYCHOTHERAPY

5.1 Introduction

The question of how psychotherapy process measures differentiate between various types of psychotherapy is key both to the validation of these measures and to understanding how psychotherapies really differ in their approach to and capacity for creating change. This question has historically been under-studied in research projects that compare the efficacy of major modes of treatment, but has gained more attention as the Dodo Bird finding (see Chapter 2) has recurred and the emphasis has shifted to process-outcome research. In this chapter, we will begin by reviewing some theoretical predictions about how psychoanalysis and psychodynamic psychotherapy differ, followed by empirical findings on how psychodynamic psychotherapies differ from psychoanalysis, as well as from non-psychodynamic therapies. Next we will introduce some evidence for how psychoanalytic process variables change over time in psychoanalysis and in psychotherapy. Based on these findings, we will propose a series of hypotheses and test them with the YAWRS applied to the Young Adult sample. Given the small size of our sample and the preliminary nature of our investigation into the YAWRS, hypothesis testing will be largely exploratory, with an eye to validating newly introduced subscales of YAWRS items. Robustness of these findings and explanations for the failure to confirm individual hypotheses, derived from existing process-outcome literature, will be discussed.

5.1.1 Theoretical differences between psychoanalysis and psychodynamic psychotherapy

The psychoanalytic literature spanning the last 50 years contains a large and often controversial debate about the theoretical distinctions between psychoanalysis and psychodynamic psychotherapy (Allison, 1994; Kernberg, 1999; Rangell, 1954; Tyson & Morris, 1992; Vaughn & Roose, 1995; Wallerstein, 1995). A full review of this dialogue is

beyond the scope of this chapter, but some of the general principles are relevant to our review of the empirical literature and proposal of hypotheses. Two questions underlie much of the discussion: (1) is the distinction between psychoanalysis and psychodynamic psychotherapy best defined by a difference in (a) therapeutic techniques, (b) therapeutic process, or (c) therapeutic goals, and (2) does the distinction lie along a continuum, in which there is no sharp boundary between the two (i.e., the quantitative theory), or is there a distinct qualitative difference between the two treatments. Rangell and Gill (1954), in a widely-cited panel on this question, agree that the appearance and subsequent analysis and resolution of the transference neurosis represents the distinguishing feature of psychoanalysis, though Gill emphasizes the difference in the resulting difference in technique, while Rangell emphasizes the effect on process.

Tyson and Morris (1992), revisiting this debate in a contemporary panel, point out that much of the disagreement still exists. According to some (e.g., Weinschel), the principle difference lies in the goals (i.e., psychoanalysis aims for substantial structural change, while psychotherapy aims only for symptom reduction). For others (e.g., Levy) the practical constraints of psychodynamic psychotherapy limit the extent of controlled regression and the freedom with which the infantile and transference neuroses can develop. Tyson points out that the distinction between psychoanalysis and psychodynamic psychotherapy becomes even more confused by the fact that psychodynamic psychotherapy may be fundamentally different when done by an analyst and a non-analyst. He points out that the most important, and still unanswered, questions may be: (1) is transference neurosis necessary for analysis, and (2) can transference neurosis be established and resolved in a psychotherapy?

Several contemporary psychoanalysts have discussed the distinction of psychoanalysis and psychodynamic psychotherapy in the broader context of the many different types of psychoanalysis and psychodynamic therapy now in existence, and how

they are best taught, prescribed, and conducted (Allison, 1994; Kernberg, 1999; Vaughn & Roose, 1995; Wallerstein, 1995). Wallerstein and Allison argue for seeing a continuum of techniques and suggest that only well-designed, psychoanalytically sophisticated empirical research will help us to see how technique, process, goals, and outcome are related, leading ultimately to deciding which treatments are indicated for which patients. Based on his findings in the Menninger PRP (see Chapter 1), Wallerstein suggests that the distinction between expressive and supportive techniques and psychotherapies may ultimately be the more important one (Wallerstein & DeWitt, 1997). Kernberg (1999) presents the clearest contemporary differentiation between psychoanalysis and psychodynamic psychotherapy, along with implications for current practice, training and research. He states that even though the stated objectives of these treatment are different (fundamental structural change vs. partial reorganization of structure in the context of symptomatic change), the considerable overlap in treatment effects, particularly in the case of patients with severe personality disorders, makes this distinction insufficient for distinguishing the treatments. He believes that the two are best differentiated quantitatively by the extent of three essential technical features of psychoanalysis: interpretation, transference analysis, and neutrality. Furthermore, these differences in “strategy” (to be distinguished from “tactics,” which do not fundamentally differ), are observable only over time, and may not appear in individual sessions.

Vaughn and Roose (1995) do not directly address the comparison of psychoanalysis and psychodynamic psychotherapy but review current theories and measures relevant to the question of analytic process, that is, what makes psychoanalysis distinct from other forms of therapy. They suggest four essential components that, with proper techniques, could be measured and used to quantify the difference between therapies: free association, resistance, interpretation, and working through. Reasonably reliable and valid measures already exist for the first three of these components, however,

there is still no practical way to quantify the extent of working through in an analytic session or treatment. More theoretical work may be needed in this area to make it possible for such a measure to be designed and effectively used.

5.1.2 Empirical evidence of differences between psychoanalysis and psychodynamic psychotherapy

The empirical evidence for process differences between psychoanalysis and psychodynamic psychotherapy is surprisingly small, given the number of studies that have looked at other issues regarding psychoanalytic process and difference in outcome. With the exception of only the Menninger PRP, all of the studies published to date that have compared psychoanalysis and psychodynamic psychotherapy have studied the difference in treatment outcome without characterizing the underlying differences in process (the outcome results from these studies will be reviewed in Chapter 6) (Bachrach, 1993; Fonagy & Target, 1996; Hamburg et al., 1967; Heinicke & Ramsey-Klee, 1986; Kordy et al., 1983; Rudolf, 1991; Sandell et al., 2000). Methodological difficulties in studying process and the relatively recent development of practical and analytically-informed process measures have been cited as the most likely cause for the lack of adequate differentiation between treatments.

The Menninger PRP (Wallerstein, 1986; 1995) has been the only large scale psychoanalytic outcome study to use therapist report measures and review of therapist process notes to quantify the differences between psychoanalysis and psychodynamic psychotherapy. Wallerstein reports that once-weekly psychotherapy included a higher frequency of supportive techniques than did psychoanalysis, though psychoanalysis included more supportive elements than had been expected. Kordy and colleagues (1983) did not study process per se, but asked patients about their level of satisfaction in treatment. He found that although analytic patients were noted to have higher rates of success in treatment by an independent observer (72% as compared to 50% for

psychotherapy), rates of satisfaction in analytic patients, as reported at termination were significantly lower (16% “very satisfied” as opposed to 52% for psychotherapy).

5.1.3 Empirical evidence of differences between psychoanalytic and non-psychodynamic psychotherapy

A significantly more productive research area has been the comparison of process in psychoanalytic and non-psychodynamic psychotherapy. In an era of managed care and short-term psychotherapy, this has proven to be a useful and more practical line of process research. Given the theoretical argument that differences in process between psychoanalysis and psychodynamic psychotherapy are largely quantitative, the findings of the analytic versus non-analytic psychotherapy process research are germane to the questions posed in this chapter. Several general measures of therapeutic process have validated their scales by finding expected differences between analytic and non-analytic treatments. Using the VPPS, Gomes-Schwartz (1978) confirmed that as compared with Rogerian therapists, psychodynamic therapists use more exploratory techniques and are less warm and friendly, while non-professional therapists give more advice and engage in more informal conversation.

Three studies employing the Jones PQS have demonstrated large and predictable differences between process in psychodynamic and cognitive-behavioural therapy (Ablon & Jones, 1998, 1999; Jones & Pulos, 1993). In the first, Jones and Pulos analysed ratings of the 100 item PQS using verbatim transcripts from 186 treatment sessions taken from 30 brief psychodynamic and 32 cognitive-behavioural therapies. The items most and least characteristic of the two types of therapy were almost entirely different, and of the 100 items 57 statistically differentiated between the two treatments. As expected, the vast majority of these items fell into the categories of therapist technique and therapist stance (the 100 Q-sort items are more or less evenly distributed between patient utterances, patient state of mind, therapist technique, therapist stance, and patient-therapist interaction). Dynamic therapists were more likely to encourage or facilitate patient

speech, identify recurrent patterns, point out the use of defences, draw attention to thoughts and feelings regarded by the patient as unacceptable or not clearly in awareness, and promote experiences of affect. CBT therapists were more likely to give direct advice or guidance, suggest activities, explain the behaviour of people in the patient's life and encourage new ways of behaving with them, give specific focus to sessions, attend to cognitive beliefs, and act to avoid or suppress patient's disturbing feelings or ideas. In terms of therapist stance, dynamic therapists were found to be more distant and formal, more neutral, more empathic, and more likely to correctly perceive emotion and interaction. CBT therapists were more controlling, didactic, tactless, condescending, and allowed their own emotions to intrude into the session, but they were also more approving, reassuring, and accommodating. Patients in dynamic therapy expressed more anger and aggression and struggled harder to control their feelings. Patients in CBT were more compliant, controlling, ambivalent, and critical. Finally, in terms of session content, dynamic sessions included more discussion of the therapy relationship and revealed a greater achievement of self-understanding on the part of the patient.

Subsequent factor analysis of the Q-sort items for all 186 sessions revealed four significant factors, labeled according to their content (1) psychodynamic technique, (2) CBT technique, (3) patient resistance, and (4) patient negative affect. The psychodynamic technique factor empirically crystallized 10 items that were found most distinctively characteristic of dynamic sessions (in decreasing order of specificity to this factor): therapist emphasizes feelings to help patient experience them more deeply, therapist is neutral, therapist interprets unconscious content, therapist points out defences, patient links feelings or perceptions to the past, therapist draws attention to feelings regarded by the patient as unacceptable, memories or reconstruction of childhood are discussed, therapist draws connections between therapeutic and other relationships, patient's in-session behaviour is reformulated, and therapist identifies a recurrent theme.

Jones and colleagues duplicated these findings in a clinically relevant way by asking expert panels of 11 experienced psychodynamic and 10 cognitive-behavioural therapists to rate the 100 Q-sort items in terms of how characteristic they would be of an “ideally conducted” treatment session of their modality (Ablon & Jones, 1998). The overlap between the most characteristic items from this method, and those identified in the earlier study was significant, with the addition of 11 new items (dreams are discussed, therapist is empathic, therapist is nonjudgmental, patient achieves a new insight, therapy relationship is discussed, therapist communicates clearly, sex is discussed, self-image is discussed, therapist facilitates patient’s speech, therapist comments on changes in mood, therapist focuses on guilt). In these two studies, as well as a third one in which transcripts from interpersonal and cognitive-behavioural therapy sessions part of the NIMH Collaborative Depression Study were scored on the PQS (Ablon & Jones, 1999), the association between various factors and treatment outcome was studied. In all three studies “psychodynamic technique” was significantly correlated with positive outcome, not only in the dynamic therapies but in the non-dynamic therapies as well (see Chapter 6 for further discussion of this finding).

Early general measures of therapist interventions, by Hill (1986), Stiles (1986), and Piper (1987) were able to differentiate partially between different modes of psychotherapy, but have been unhelpful in predicting psychotherapy outcome. This led to measures that focused on theoretically-specific interventions, such as Gaston and Ring’s ITS (Gaston & Ring, 1992) and Cooper and Bond’s PIRS (Milbrath et al., 1999). Preliminary research using the ITS has shown higher ratings of exploratory technique, and lower ratings of supportive and work-enhancing strategies in dynamic, as compared to cognitive, therapy (Gaston & Ring, 1992).

Observer and therapist-rated measures of transference interactions and interpretations have been developed to relate these specific interventions to outcome in

psychodynamic psychotherapy and psychoanalysis. Though not all research groups have bothered to demonstrate a contrast with non-psychoanalytic treatment, this has been a useful confirmation of validity in some studies. Bøgwald and colleagues (1999) showed a significant difference between the extent of transference interpretations in two forms of manualized brief dynamic therapies, one with a focus on transference and one without.

In the recent climate of randomized comparisons and empirically validated psychotherapies (see Chapter 2), there has been an effort to adapt therapy process measures for measuring the differences between and adherence to alternative psychotherapies. One of the first and most well known of these studies, the NIMH Treatment of Depression Collaborative Research Program, utilized a measure (the Collaborative Study Psychotherapy Rating Scale, CSPRS) specifically designed for this purpose (Elkin et al., 1989; Hill, O'Grady, & Elkin, 1992). This 96-item observer-rated checklist successfully classified the therapies in the study (CBT, IPT, and placebo or medication) 95% of the time, but did not attempt to collect data on alternative interventional strategies (Ablon & Jones, 1999). Many other such instruments have been developed to measure therapeutic technique, including Orlinsky's therapist-completed Therapeutic Procedures Inventory (TPI-R, McNeilly & Howard, 1991). A factor analysis of the measure's 73 items yielded three scales that matched theoretical predictions: directive/behavioural, psychodynamic/past-focused, and affective.

5.1.4 Empirical evidence for change in psychoanalytic process over time

A small body of research has been assembled that looks specifically at the change in psychoanalytic process variables over the course of analysis. While this does not provide direct evidence for how psychoanalysis and psychodynamic psychotherapy differ on process, the change over time is relevant to examination of these patterns and to the variables relevant to these treatments. Graff & Luborsky (1977), using the 23 item therapist checklist described in Chapter 4, found in two successful analyses that therapist

reported transference increased over time while resistance decreased. Meanwhile, in two less successful treatments, transference and resistance both remained stable and were associated with one another.

Jones and colleagues have applied the PQS to single psychoanalytic cases, in order to test specific psychoanalytic hypotheses. In one study, time-series analysis was used to demonstrate that particular types of analyst interventions (interpretation of defences, identifying a recurrent theme in material, and discussion of dream or fantasy material) led to an increase in a patient's associative freedom (Jones & Windholz, 1990; Spence et al., 1993). In another, psychodynamic technique was found to lead to free association (measured by the co-occurrence of specific speech patterns), which in turn was linked to awareness of guilt and symptom improvement (Pole & Jones, 1998).

Several studies of psychoanalysis have looked specifically at the transference neurosis, testing the hypothesis, put forth particularly in the early theoretical literature, that the resolution of this neurosis is essential to good psychoanalytic outcome (Norman et al., 1976; Oremland et al., 1975; Schlessinger & Robbins, 1974). The surprising consensus has been that, even in successful treatments, the neurosis is not fully obliterated and resurfaces easily in interviews years later. A separate but equally informative finding in several large scale studies of psychoanalysis is that greater pathology in the patient is revealed during the treatment than was detected at the outset (Erle, 1979; Kantrowitz, 1993; Sandell, 2000; Wallerstein, 1986).

5.1.6 Empirical evidence for change in psychotherapeutic process over time

Finally, and perhaps least relevantly for a study of psychoanalytic process, there have been a number of findings about the course of dynamic process variables during the course of a psychotherapy. In a study of 72 cases, using Orlinksy's TPI-R, affective scores were observed to rise steadily, while psychodynamic/past-focused scores peaked at mid-treatment and then declined (McNeilly & Howard, 1991). Other work has

elucidated complex interactions between therapist interventions and therapy process. Two studies of a long-term therapy (Jones, Ghannam, Nigg, & Dyer, 1993; Jones & Price, 1998) used sequential analysis to demonstrate that there is mutual influence between patient and therapist. In the treatment studied, the therapist began by being nonjudgmental, facilitative, and neutral, but the patient's depressive affect gradually drew him towards a more actively challenging and emotionally reactive posture, which in turn led to symptom reduction in the patient.

Russell and Trull (1986) summarize convergent findings from 33 studies (1945 through 1984) attempting to detect patterns in language oriented process: (1) less controlling therapist interventions precede client insight or self-exploration, (2) therapist interpretations of moderate depth or focused on manifest content were less likely to precede client resistance than were deeper interpretations, (3) successful topic initiation and subsequent uptake were higher in therapeutic dyads matched on role expectancies, and (4) therapists tended to follow client hostility directed at the therapist with avoidance, and novice therapists tended not to interpret client resistive statements. From this range of findings and from the diversity of language-oriented measures, it appears clear that this approach has much to offer psychotherapy research but that a great deal of work still remains in order to understand the advantages and disadvantages of different methods.

The Menninger TRIP study (Gabbard et al., 1994) concluded that transference interpretations tended to have greater impact – both positive and negative – than other interventions made with these patients and identified certain factors which predicted negative response to these interpretations: neuropsychologically based cognitive dysfunction, history of early trauma, patterns of object relations involving interpersonal distance, masochistic tendencies, and anaclitic rather than introjective psychopathology (Gabbard et al., 1994). As foreshadowed by the Menninger PRP 40 years earlier (see

Chapter 1 for summary of findings), transference interpretations were found to be more helpful in strengthening the therapeutic alliance if used after supportive interventions. Several other studies have found significant effects of transference interpretations on aspects of the therapeutic process, particularly alliance (Ackerman & Hilsenroth, 2003; Connolly, 1998; 1999; Høglend, 1993c; Piper et al., 1991).

5.1.5 Summary

Though research that specifically compares the process in psychoanalysis and psychodynamic psychotherapy studies has been limited, a number of other sources have supplied significant information in this crucial area. Theoretical writing has emphasized the quantitative and qualitative importance of basic psychoanalytic processes including transference interpretation, transference neurosis, free association, general interpretation, resistance, and working through. When process measures have been used to quantify the differences between psychoanalytic and non-psychodynamic psychotherapy, a wide range of expected differences were found, most significantly along the lines of therapist interpretations and certain types of content in the discussion (e.g., feelings, unconscious themes, defences, links to the past, transference, in-session behaviour, dreams, sex, guilt, and self-image). Negative patient views of treatment have been found to be more common in psychoanalysis. A few temporal patterns in psychoanalysis and psychodynamic psychotherapy have been identified as well, particularly an increase in transference, an increase in associative freedom, and an increase in affective themes.

Though it has not been a subject of empirical psychoanalytic research, it stands to reason from the clinical literature on psychoanalysis that the material raised in analysis will be more evocative of therapist affective response than that in psychodynamic psychotherapy. A number of analytic theorists including Melanie Klein (1948) and Herbert Rosenfeld (1965) have emphasized the extent of primitive themes emerging in psychoanalytic work, which would likely be quantitatively and perhaps qualitatively

different from that which arises in psychodynamic psychotherapy. Many clinical writers have commented on the way in which this material evokes powerful negative affects in clinicians, often falling under the title of a negative countertransference (e.g., Brenman-Pick, 1985).

5.2 Hypotheses

Based on the scattered but interesting evidence from clinical, theoretical, and empirical sources, nine hypotheses are proposed about the nature of psychoanalytic process captured by the YAWRS (described in Chapter 4). As consistent process data in this sample is only available during the first year of treatment, the hypotheses are formulated in terms of using the data from this first year to differentiate the two treatment groups. These hypotheses are designed to be validity checks on the associated constructs in the YAWRS as well as investigating answers to these basic questions in this sample.

Therapist variables

1. Extent of dynamic technique, as operationalised by Jones PQS, will be greater in psychoanalysis (Ablon & Jones, 1998; Gaston & Ring, 1992; Jones & Pulos, 1993; Milbrath et al., 1999).
2. Overall extent of interpretation will be greater in psychoanalysis (Kernberg, 1999).
3. Extent of transference interpretations will be greater in psychoanalysis (Bøgwald et al., 1999; Kernberg, 1999; McGlashan & Miller, 1982) and will increase over time in both treatments (Graff & Luborsky, 1977).
4. Extent of supportive interventions will *not* be greater in psychotherapy (Ablon & Jones, 1998; Jones & Pulos, 1993).
5. Negative countertransference will be greater in psychoanalysis (Klein, 1948; Rosenfeld, 1965).

Patient variables

6. Extent of dynamic material, as operationalised by Jones PQS, will be greater in psychoanalysis (Ablon & Jones, 1998; Jones & Pulos, 1993; Vaughn & Roose, 1995) and will increase over time in both treatments (Jones & Windholz, 1990; McNeilly & Howard, 1991; Pole & Jones, 1998; Spence et al., 1993).
7. Extent of transference themes will be greater in psychoanalysis (Kernberg, 1999; Vaughn & Roose, 1995) and increase over time in both treatments (Graff & Luborsky, 1977).
8. Extent of regression will be greater in psychoanalysis (Rangell, 1954; Tyson & Morris, 1992; Vaughn & Roose, 1995) and increase over time in psychoanalysis (Klein, 1948; Rosenfeld, 1965).
9. Patient's negative view of treatment will be greater in psychoanalysis (Kordy et al., 1983; Russell & Trull, 1986) and decrease over time (Graff & Luborsky, 1977; Russell & Trull, 1986).

5.3 Methods

The basic experimental design, including subject and assessments, has been described in Chapters 3 and 4. Data analysis was performed on the 10 Young Adult subjects for whom at least 6 months of YAWRS data was collected during the first year of analytic treatment. As was demonstrated in Chapter 3, there were no significant differences found in demographic or diagnostic variables between the 10 subjects in this analysis and the 14 subjects that were excluded.

5.3.1 Data processing

The final step in preparation of the YAWRS data for analysis entailed accommodating the empirically derived factors and scales derived to this point with the theoretical intent of the hypotheses presented in the introduction. Each theoretical

variable mentioned in one of the hypotheses was matched with the derived global factors, YAWRS factors, subsection summary scores, and individual YAWRS items, in that order of preference. The minimum number of factor analytic, subsection summary, and item scores with the greatest breadth were selected to correspond to each hypothesis scale, so as to cover the concept adequately and use empirically derived means for maximizing scale reliability and coherence.

5.3.2 *Data analysis*

The relationships of process scales to treatment intensity and outcome were tested using two data analytic methods. In the first, more descriptive method, mean values were calculated on each of the 15 scales for each subject on all available YAWRS data collected during the first year of treatment. The number of values that entered into these means varied between 6 (for two psychotherapy subjects for whom data from only half a year was available) and 42 (for a psychoanalytic subject who had attended regularly all year). The mean number of YAWRS data points per subject was 20.1 with a standard deviation of 15.5. Student's independent samples t-tests were performed to test the difference between psychoanalytic and psychotherapy cases for the scales suggested in Hypothesis 3 and between symptomatic improvers and non-improvers for the scales suggested in Hypothesis 4. Due to the exploratory nature of this analysis and the fact that individual predictions were made for each of these analyses, Bonferroni corrections were not used and the α was set at 0.05.

In the second method, an attempt was made to compare linear trends as well as mean differences between psychotherapy and psychoanalytic patients and symptomatic improvers and non-improvers on the 15 scales. Hierarchical linear modeling (HLM), a technique developed for comparing irregular data sets in a nested structure, such as time series data for a series of subjects divided into comparison groups, was selected for this purpose (Bryk & Raudenbush, 1987; 2001; 1996). The HLM/2L software procedure uses

an iterative technique to model the relationship between group membership (intensity of treatment or symptomatic improvement) and linear parameters of YAWRS scales taken over time (slope and intercept). Because the psychoanalytic and improvement groups only differ by a single subject and the HLM procedure is optimized for finding differences between groups (which in this case would be highly subject to the characteristics of that single subject), the HLM procedure was applied separately for the two grouping variables to each scale. A fixed effect estimate of the relationship between group membership and intercept or slope that was significantly different from 0 in the direction predicted was considered to be a significant finding. As above, because of the exploratory nature of these analyses and the specificity of the predictions, Bonferroni corrections were not used and the α was set at 0.05.

5.4 Results

5.4.1 *Generation of hypothesis-based scales*

Based on the results of the subsection and subsection summary factor analyses, sets of variables were chosen to correspond to the process constructs outlined in the introduction and hypotheses of this chapter. In order to maximise the empirical basis for these scales, yet cover the constructs adequately, scales were assembled with the following order of preference: (1) global factors (GLF), (2) subsection factors (FAC), (3) subsection summary scores (SSC), and (4) individual YAWRS items (ITM). The composition of the seven therapist scales is outlined in Table 5.3. The six patient scales are described in Table 5.4. Formulae for all 15 scales are given in Appendix 5.2d.

1. Jones dynamic technique

- a. Interprets feelings: interprets sadness [PC2oi, SSC], primitive emotional stance [PC3oi, SSC], and anger [PC3oi, SSC]
- b. Interprets unconscious: interprets general unconscious [UAoi, SSC], unconscious reactions to aggression [UBoi, SSC], unconscious content related to sexuality [UCoi, SSC], unconscious content related to self and self-esteem [UDoi, SSC], and unconscious content related to body [UEoi, SSC]
- c. Interprets defences: analyst comments on defensive manoeuvres [UG15, ITM], patient's resistance [UG16, ITM], patient's relationship to analytic process [UG17, ITM], interprets defences: projection, reaction formation, denial [PDF1i, FAC], isolation [PDF2i, FAC], splitting [PDF3i, FAC], regression, projective identification [PDF4i, FAC], and externalisation [PDF5i, FAC].
- d. Interprets links to past: interprets historical material [MCOi, SSC], comments about patient's experiences of little child or baby within [UG12, ITM]
- e. Interprets transference: see item 3 below.
- f. Interprets in-session behaviour: interprets resistance [GAoi, SSC], general attitude to analysis [GBoi, SSC], time keeping [GCoi, SSC], missed sessions [GDoi, SSC], patient's behaviour in session [GEoi, SSC], patient's aggression and sexuality in the analysis [GGoi, SSC].
- g. Interprets clear themes: interprets clear theme to material [GF02i, ITM], completeness and quality of analyst's understanding of patient's material [J3, ITM], analyst comments on links between behaviour outside the session to current treatment material [UG13, ITM].
- h. Interprets dreams: interprets dreams [GF05i, ITM], daydreams and fantasies [GF06i, ITM], associations to dreams and fantasies [GF07i, ITM], meaning or significance of thoughts and fantasies [GF08i, ITM].
- i. Interprets sex: interprets sexual relations [MFOi, SSC], sexuality [MKoi, SSC], unconscious themes related to sexuality [UCoi, SSC].
- j. Interprets self-image: interprets manifest themes related to body [MAoi, SSC] and self-esteem [MBoi, SSC], interprets unconscious themes related to self and self-esteem [UDoi, SSC] and unconscious content related to body [UEoi, SSC].
- k. Interprets guilt: interprets manifest guilt, punishment, and conflicted loyalty themes related to historical material [MC04ci, ITM], family [MD04bi, MD04hi, ITM], friends [ME04ci, ME04ji, ITM], sexual relations [MF04di, MF04ji, ITM], sexuality [MK06di, ITM], interprets guilt concerning sexual wishes [PC02ci, ITM], aggressive wishes and actions [PC02di, ITM], striking absence of guilt [PC03fi, ITM], fear of being humiliated [PC03ii, ITM], intense shame [PC03ji, ITM], and overwhelming guilt in relation to analyst [UF04bi, ITM].

Table 5.3. Composition of therapist-based hypothesis scales (continued).

-
2. General interpretation: all interpretation of general, manifest, preconscious, and unconscious content [GAoi-GHoi, MAoi-MLoi, PA1oi-PDF5i, UAoi-UF4i, UGf1-UGf4, FAC and SSC].
 3. Transference interpretation: interprets general transference [PA1oi, SSC], positive wishes towards analyst [PA2oi, SSC], transference with anxiety [PA3oi, SSC], transference with competitive and aggressive themes [PA4oi, SSC], transference with resentment [PA5oi, SSC], and primitive transference [PA6oi, SSC], interprets change in transference across week [PBoi, SSC], analyst comments on connections between therapeutic relationship and past relationships [UG09, ITM], connections between therapeutic relationship and present relationships [UG10, ITM], and displacing feelings from the analyst to outside figure [UG11, ITM].
 4. Supportive interventions: supportive intervention [UGf3, FAC]
 5. Therapist's positive view of treatment: analyst feels patient was helped [UF3p, SSC], analyst feels loving/empathic [UHF3o, FAC], analyst's judgement of "quality of week" [Javg, SSC], -analyst feels attacked/disgusted [UHF1o, FAC], -inadequate/confused [UHF2o, FAC], and -bored/cutoff [UHF4o, FAC], -patient is bullying [UIf1o, FAC], -narcissistic [UIf2o, FAC], and -analyst is rejected [UIf3o, FAC].
-

Table 5.3. Composition of therapist-based hypothesis scales.

-
1. Jones dynamic material
 - a. Feelings: sadness [PC2op, SSC], primitive emotional stance [PC3op, SSC], anger [PC3op, SSC]
 - b. Unconscious content: general [UAop, SSC], reactions to aggression [UBop, SSC], related to sexuality [UCop, SSC], related to self and self-esteem [UDop, SSC], related to body [UEop, SSC]
 - c. Defences: projection, reaction formation, denial [PDF1p, FAC], isolation [PDF2p, FAC], splitting [PDF3p, FAC], regression, projective identification [PDF4p, FAC], externalisation [PDF5p, FAC].
 - d. Links to past: historical material [MCOp, SSC]
 - e. Transference: see item 2 below.
 - f. In-session behaviour: resistance [GAop, SSC], general attitude to analysis [GBop, SSC], time keeping [GCop, SSC], missed sessions [GDop, SSC], bad behaviour in session [GEOp, SSC], patient's aggression and sexuality in the analysis [GGop, SSC].
 - g. Clear themes: clear theme to material [GF02p, ITM].
 - h. Dreams: dreams [GF05p, ITM], daydreams and fantasies [GF06p, ITM], associations to dreams and fantasies [GF07p, ITM], meaning or significance of thoughts and fantasies [GF08p, ITM].
 - i. Sex: sexual relations [MFop, SSC], sexuality [MKop, SSC], unconscious themes related to sexuality [UCop, SSC].
-

Table 5.4. Composition of patient-based hypothesis scales (continued).

-
- j. Self-image: manifest themes related to body [MAop, SSC] and self-esteem [MBop, SSC], unconscious themes related to self and self-esteem [UDop, SSC] and unconscious content related to body [UEop, SSC].
 - k. Guilt: manifest guilt, punishment, and conflicted loyalty themes related to historical material [MC04cp, ITM], family [MD04bp, MDo4hp, ITM], friends [ME04cp, ME04jp, ITM], sexual relations [MF04dp, MF04jp, ITM], sexuality [MK06dp, ITM], interprets guilt concerning sexual wishes [PC02cp, ITM], aggressive wishes and actions [PC02dp, ITM], striking absence of guilt [PC03fp, ITM], fear of being humiliated [PC03ip, ITM], intense shame [PCO3jp, ITM], and overwhelming guilt in relation to analyst [UF04bp, ITM].
2. Transference themes: general [PA1op, SSC], positive wishes towards analyst [PA2op, SSC], transference with anxiety [PA3op, SSC], transference with competitive and aggressive themes [PA4op, SSC], transference with resentment [PA5op, SSC], primitive transference [PA6op, SSC], change in transference across week [PBop, SSC].
 3. Regression: aggression and sexuality in the analysis [GGop, SSC], immaturity of patient's mental functioning [GHop, SSC].
 4. Patient's positive view of treatment: positive wishes towards analyst [PA2op, SSC], patient feels helped [UF1p, SSC], –negative attitude [GBop, SSC], –transference with anxiety [PA3op, SSC], –transference with competitive and aggressive themes [PA4op, SSC], –transference with resentment [PA5op, SSC], –primitive transference [PA6op, SSC], –patient reports not feeling helped [UF2p, SSC], –analyst feels patient reacts negatively [UF4p, SSC].
-

Table 5.4. Composition of patient-based hypothesis scales.

5.4.2 Association between YAWRS scales and intensity of treatment

Results of the mean comparisons of patients in psychoanalysis and psychodynamic psychotherapy on the nine variables outlined in the hypotheses are shown in Table 5.5. Using this method, all eight predicted differences were observed in the expected direction, and five reached statistical significance. In psychoanalysis, as compared to psychodynamic psychotherapy, therapists reported using more general interpretation, more transference interpretation, and had a less positive view of the treatment. In addition, therapists reported that patients shared more dynamic material (based on Jones's definition of dynamic technique) and spoke more about the transference. As expected, no significant difference was found in therapist report of supportive interventions between patients in psychoanalysis and psychodynamic psychotherapy.

There was a non-significant trend ($0.05 < p < 0.1$) toward a higher level of therapist reported dynamic technique and a lower level of patient positive view of the treatment in psychoanalytic patients. The difference in therapist-reported patient regression did not reach the level of a trend.

	Mean (SE)		Student's T-test (df=8)
	Psychoanalysis (n=4)	Psychodynamic psychotx (n=6)	
Therapist variables			
(a) Jones dynamic technique	0.86 (0.11)	0.61 (0.08)	1.88 [†]
(b) General interpretation	0.72 (0.11)	0.41 (0.09)	2.29*
(c) Transference interpretation	0.90 (0.07)	0.54 (0.11)	2.43*
(d) Supportive interventions	0.46 (0.24)	0.45 (0.21)	0.04
(e) Therapist's positive view of treatment	1.80 (0.16)	2.28 (0.05)	-3.33**
Patient variables			
(f) Jones dynamic material	0.88 (0.09)	0.60 (0.07)	2.49*
(g) Transference themes	0.71 (0.08)	0.36 (0.11)	2.37*
(h) Regression	0.76 (0.16)	0.41 (0.15)	1.52
(i) Patient's positive view of treatment	1.97 (0.05)	2.20 (0.09)	-1.89 [†]

Table 5.5. Results of T-test analysis: Psychoanalysis versus psychodynamic psychotherapy.

[†] $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

Results of the HLM analysis are reported in Table 5.6. Significant mean differences, as reflected in an estimated treatment intensity intercept significantly different from 0, were found in the expected direction for Jones dynamic technique and Jones dynamic material. A trend was found, also in the predicted direction, for general interpretation. No differences were found for transference interpretation, supportive interventions, patient regression, or therapist and patient's view of treatment. For all nine scales tested, no significant linear trends were found either in the sample as a whole or in relation to treatment intensity. Thus predictions that Jones dynamic techniques, extent of

interpretation, extent of transference interpretation, Jones dynamic material, and extent of transference themes would increase over time were not supported by this analysis.

	Fixed effect estimate (SE)		
	Treatment intensity → intercept	Baseline slope	Treatment intensity → slope
Therapist variables			
(a) Jones dynamic technique	0.35 (0.14)*	0.003 (0.002)	-0.003 (0.002)
(b) General interpretation	0.34 (0.16) [†]	0.001 (0.002)	-0.001 (0.002)
(c) Transference interpretation	0.39 (0.24)	0.000 (0.003)	-0.001 (0.004)
(d) Supportive interventions	-0.05 (0.27)	0.001 (0.004)	0.002 (0.005)
(e) Therapist's positive view of treatment	-0.43 (0.23)	0.000 (0.004)	-0.001 (0.006)
Patient variables			
(f) Jones dynamic material	0.33 (0.11)*	-0.001 (0.002)	-0.002 (0.003)
(g) Transference themes	0.41 (0.19) [†]	0.002 (0.003)	-0.003 (0.004)
(h) Regression	0.23 (0.30)	-0.002 (0.003)	0.005 (0.004)
(i) Patient's positive view of treatment	-0.15 (0.22)	0.003 (0.004)	-0.003 (0.005)

Table 5.6. Results of HLM analysis of Hypothesis 3: Psychoanalysis versus psychodynamic psychotherapy.

[†] p < 0.1, * p < 0.05, ** p < 0.01

5.5 Discussion

5.5.1 Findings

Results from the comparison of psychoanalysis and psychodynamic psychotherapy show that the YAWRS consistently distinguishes these two treatments on the basis of therapist report in the first year. Statistical significance via T-test on five of the eight predicted differences, with suggestive trends on the three other variables, is an encouraging sign of the validity of the YAWRS as a relevant measure of psychoanalytic process and is an exploratory start at answers to theoretical questions posed in the introduction. These results suggest that despite the commonly described similarity in

therapist “tactics” between psychoanalysis and psychodynamic psychotherapy (Kernberg, 1999; McGlashan & Miller, 1982) there are demonstrable quantitative differences in a process measure designed to capture relevant components of analytic technique. This argues against one criticism of Jones’s measure, that his constructs represent “common factors” and are not specific to either psychoanalysis or psychodynamic psychotherapy (Markowitz, personal communication, August 2003).

The greatest weakness of this finding is that the YAWRS is only therapist report and as the therapist is not blind to the type of treatment he is rating, the scores may reflect more his theoretical bias more than a real difference in treatment process. On the other hand, evidence that therapists are not unusually biased in their ratings of patient variables (Weiss et al., 1996) and the concrete nature of the YAWRS items lends support to the meaningfulness of these results. At the very least, the YAWRS scales appear to capture distinctions between psychoanalysis and psychodynamic psychotherapy shared by the clinicians in this study. More speculatively, the results provide early evidence for our contention that the therapist is a sophisticated and uniquely positioned informant who by aggregating data from multiple sessions can usefully capture important clinical differences in therapeutic modalities.

The most significant validity check emerged from the use of the Jones PQS to design “dynamic technique” and “dynamic material” aggregate process variables. These variables included items with a wide range of analytic phenomena and were well-validated by Jones and colleagues on prior comparisons of analytic and non-analytic therapy (Ablon & Jones, 1998, 1999; Jones & Pulos, 1993). Findings that the dynamic material variable was higher in psychoanalysis using both T-test and HLM analyses, and that the dynamic technique variable was higher in psychoanalysis using the HLM analysis support the claim that, by therapist report, there are significant quantitative differences between the two treatments along the same continua that distinguish psychoanalytic from non-

psychodynamic therapy. This not only supports the validity of the YAWRS, indicating that it shares some attributes of the PQS, but also suggests that there is broad truth in the theoretical claim that the therapies are distinguishable along these lines. These results also build on the findings from Gaston and Ring's ITS (1992) and Cooper and Bond's PIRS (Milbrath et al., 1999) that "exploratory techniques" are rated higher in a more dynamic therapy.

The more specific findings that the YAWRS general interpretation, transference interpretation, and transference theme scales all capture significant quantitative differences between treatment groups, support the theoretical literature in this area. Kernberg's (1999) emphasis on interpretation and transference interpretation as essential to psychoanalytic process is consistent with these results. This is the first step in support of his notion that the difference between psychoanalysis and psychodynamic psychotherapy can be captured by *quantitative* differences in technique and material that lead to *qualitative* outcome differences. These data also support the suggestions of Vaughn and Roose (1995) and McGlashan (1982) that transference-related content will be most identifiable with psychoanalysis.

Of all the distinctions proposed by theoreticians about the difference between psychoanalytic and psychotherapy process, transference interpretations are the most concrete and consistently cited example (Kernberg, 1999; McGlashan & Miller, 1982; Tyson & Morris, 1992). Even authors who emphasize the quantitative differences between the treatments, note that there may be a qualitative distinction between the role of this central process (McGlashan & Miller, 1982). Therefore it is reasonable that the difference on this variable be significant via T-test, though somewhat surprising that it did not occur by HLM analysis as well. This may be related to the fact that the therapists performing the therapy were themselves all analysts and strong believers in the role of the use of transference. Thus the sample is not representative of all psychodynamic

psychotherapy. Also, given the fact that all process is therapist-reported, there may have been a bias towards emphasizing transference interpretations, secondary to their own value judgements that this is the most mutative technique. On the other hand, there is evidence from Bøgwald's work (1999) that the prevalence of transference interpretations, measured by a third party, reflects what therapists are instructed to do when they use a manualized treatment. It is likely that the therapists in this study believed in the use of transference interpretations and used them in psychotherapy almost as much as they did in psychoanalysis.

While the lack of a significant statistical finding in a study with such small sample size and power is not itself a significant finding, the lack of any trend distinguishing the extent of supportive interventions in psychoanalysis and psychodynamic psychotherapy is interesting and confirms expectations from some important theoretical and empirical work. In the classic literature, but to some extent even in contemporary psychoanalytic writing, supportive work has been considered "non-analytic" and would be expected to be less prevalent in psychoanalysis than in a "less intensive" psychodynamic psychotherapy (Rangell, 1954; Tyson & Morris, 1992). Even Kernberg, who is a longtime advocate of the need for supportive elements in a psychodynamic psychotherapy for patients with severe personality disorders, sees less intense treatment as a setting where more supportive work can be done (Kernberg, 1999).

However, the work of Jones, Ablon, and Pulos (Ablon & Jones, 1998, 1999; Jones & Pulos, 1993) that suggested supportive elements are not a distinguishing factor between analytic and non-psychoanalytic forms of psychotherapy and the lack of emphasis in the theoretical literature on this as a distinction between psychoanalysis and psychodynamic psychotherapy, led us to the hypothesis, consistent with the data, that supportive techniques are equally represented in the two therapies. In addition to the fact that this is only a negative finding, the supportive intervention scale is based on a single

factor (representing four YAWRS items), likely making it less reliable and valid than the other scales. Therefore it is a highly preliminary result that requires more investigation. The high rates of depression, anxiety, and personality pathology in the sample, as described in Chapter 3, is significant for this finding. Both by Kernberg's theoretical recommendations, and evidence from the Menninger PRP and TRIP studies, patients with this pathology, regardless of the type of therapy they are in, require extensive supportive interventions if they are to form an alliance and benefit from interpretations (Gabbard et al., 1994; Wallerstein, 1986). This points to a major advantage of this study, namely that patients were assigned to the treatment groups sequentially, and not based on preference or indication. In studies where assignment is based on therapist or patient preference, it is likely that even subtle differences in pathology may result in significant differences in extent of supportive style, more tied to the patient than to the therapeutic modality.

Hypotheses regarding negative transference and countertransference were most speculative, but the findings are nonetheless important and interesting. There was partial support for the findings of Kordy (1983) and Russell (1986) that patients have lower satisfaction and greater negative transference in the more intense treatment. Given that the difference in transference interpretation was not significant on HLM, it is understandable that negative transference is also not significant by this means, as there is evidence that transference interpretations (particularly when not paired with supportive interventions) are associated with negative transference reactions (Gabbard et al., 1994; Høglend, 1993c; Piper et al., 1991).

The partial finding in this study that therapist negative countertransference is greater in psychoanalysis than in psychotherapy is the first empirical finding of its type and consistent with theoretical work about the primitive nature of analytic material (Klein, 1937/1984; Rosenfeld, 1971). This is particularly striking given that the analysts

rating these scales were no doubt aware of the countertransferential components of their feelings and described them in concrete negative terms, not just as countertransference (though one could argue that they also knew that awareness of the countertransference does not change the truth of the underlying feelings). The finding may be related to the level of personality pathology in these patients and relies on the fact that the analytic and psychotherapy samples did not differ significantly at baseline and were not assigned to treatment on the basis of their pathology.

Despite the number of significant findings outlined above, several expected differences, particularly on the HLM analysis, did not emerge. Regression, long theorized to be a central process in psychoanalysis and closely linked to the intensity of treatment (and sometimes used as a justification for why such intensity is necessary) did not appear significantly different in the two treatment groups. One may conclude either that the measure was insensitive to this difference (perhaps because analysts overestimate the amount of regression in their psychotherapy patients because they are biased in their belief that it should occur) or that in a sample such as this, regression is almost as important a process in psychotherapy patients as it is in analytic patients. Lack of power and sample size are other possible explanations. Finally, we recall Kernberg (1999) pointing out that objectives and outcome should never be used to differentiate psychoanalysis and psychodynamic psychotherapy, because when each is done properly with the appropriate patient population (i.e., personality disordered patients in the case of psychotherapy), similar structural changes occur.

It is also important to note from the literature, that a great deal of variation in therapist process and technique has been observed even among therapists of the same school, while therapists of different schools can be quite similar, making it difficult to draw general conclusions (Ablon & Jones, 1998; 1999; Najavits, 2001; Smith et al., 1980). Sashin and colleagues (1975) reported large variations in psychoanalytic technique as part

of the Boston Psychoanalytic Institute outcome study. Language-based measures, such as Bucci's RA or CRA (1997a; 1998) or Spence's CORtrans and SEPtrans (Mayes & Spence, 1994; Spence, 1998), have not even been applied to the distinction, in part because the variation between any two sessions, even of the same type of treatment and within the same patient, is too large (Canfield et al., 1991; Mergenthaler & Kächele, 1996; Russell & Trull, 1986). Similarly, affective measures have been applied only to small samples of psychodynamic psychotherapy, and have not yet been used to look at differences between modalities. Once the methods have become more standard, this will be a key means for understanding how forms of therapy differ.

It was somewhat disappointing that the HLM analysis failed to record any significant linear time trends in process variables as part of this analysis, either in across groups or related to treatment intensity. In contrast with the stated hypotheses, extent of transference interpretation, dynamic material, extent of transference themes, regression, and patient's negative view of the treatment did not appear to increase over time. Methodologically, this may have resulted from low power and relative insensitivity of the HLM technique to linear trends given the small number of subjects. Alternatively, one year may be a poor time scale for observing meaningful linear trends. Kernberg (1999) points out that psychoanalytic and psychotherapeutic sessions should not be distinguishable on the session time scale, requiring longer periods of observation. Scales associated with dynamic process appear to be consistently higher in psychoanalysis and fluctuations over the year are more related to weekly or monthly changes in the analytic process than in a consistent pattern across the year.

5.5.2 Study weaknesses

The difficulties with this study are related to limitations in the measure itself, as reviewed in Chapter 4, as well as to problems with how the YAWRS was applied to this sample. First and foremost, the YAWRS is limited by the fact that it captures only the

therapist report of process. Particularly with the questions raised in this chapter, which are discussed and thought about by analysts widely, it is unclear to what extent results reflect the therapist's own theoretical and clinical views versus objective processes in the treatment. On the other hand, the specific nature of the items in the YAWRS and the use of numerous items averaged in clinically relevant ways makes it more likely that therapists report accurately. There is some evidence that therapists are less biased reporters of analytic outcome and process than has sometimes been suggested (Weiss et al., 1996). The fact that the results of this study are quite similar to those in the observer-rated literature, support this argument.

The second most significant problem in this study is the small sample size. A comparison between four psychoanalytic and six psychotherapy subjects has relatively little generalizability to the analytic and psychotherapy patient populations at large, and serves more as a first exploratory step. The methodological demands of an extensive review of process, as was performed in this study, can require compromises in other areas, particularly in sample size. A number of psychotherapy researchers, however, have convincingly argued that in the exploratory phase of research the advantages of the exhaustive and systematic process work that can be done with one or a few cases outweigh the disadvantages of small sample size (Greenberg, 1991; Hilliard, 1993; Jones, 1993a; 1990; Kächele, 1992). Several interesting case studies of psychoanalytic process have been published which propose interesting hypotheses and test them in a detailed way on one or a few subjects (Jones, Ghannam et al., 1993; 1991; Moran & Fonagy, 1987; 1998).

One inevitable consequence of any study comparing small numbers of subjects is that confounding variables which differ in the two groups may be as equally likely as group membership to explain the observed differences. In the case of this study, as shown in Chapter 3, the psychoanalytic and psychotherapy groups do not differ

significantly on any major baseline variables. However, they are highly distinct in two important ways: the length of the treatment and the rates of symptomatic improvement. Subjects in the psychoanalytic group stayed in treatment significantly longer and had a much higher rate of symptomatic improvement than did the subjects in psychotherapy. It is not possible from the present data to establish whether these are causally related in any way. However, as the YAWRS data were taken only from the first year of treatment, the measure was likely able, in most cases, to capture aspects of process prior to significant change in symptomatology or decision to end treatment. At the same time, restriction to the first year means that the differences captured do not reflect aspects of the process that develop later. Contemporary psychoanalytic clinical literature stresses how process shifts over time and the first year is likely only to reflect the introductory phase of treatment, neglecting mid-phase and termination. The natural course of psychotherapy process has been less well characterized and is probably enormously variable.

As with any treatment study, the process data here reflect directly only on the type of analysis and psychodynamic psychotherapy being performed by this relatively small group of analysts. This problem is mitigated by the experience level of the analysts in the study and their stature in the Contemporary Freudian group of the British Psychoanalytic society. The group met weekly for years to discuss their clinical and theoretical work, suggesting that there was some generalizability to it, and numerous articles and books were published by members of the group. Nevertheless, the treatment they applied differs in many ways from that of other analysts and psychotherapists around the world. Kernberg (1999) strongly suggests that proper psychodynamic psychotherapy can only be done with two sessions per week and once a week psychotherapy, no matter how rigorous the clinician tries to be, reverts to a mainly supportive treatment (albeit psychoanalytically informed). To relate the theoretical and clinical work he has done on distinguishing various types of therapy, to the current work, one might do better to see

the psychotherapy in this study as fitting somewhere between his psychoanalytic and supportive psychotherapies. The lack of larger quantitative differences on dynamic themes, interpretations, and transference interpretations, and the complete lack of difference in supportive techniques is more consistent with the idea that psychodynamic psychotherapy, as studied here, is possible once a week.

Although the therapists in this study were intended to take on one analytic and one psychotherapy patient apiece, due to the smaller numbers of subjects in the YAWRS analysis, only one therapist was represented in both intensive and non-intensive groups in this analysis (subjects M and R have the same treater). This has both advantages and disadvantages in interpreting the above results. Finding significant difference in the ratings of different therapists suggests that the results do not just indicate the theoretical distinctions made independently by a few therapists who are representing what they see and do differently in two modes of treatment, but rather some consensus amongst a group on how they differ. On the other hand, the results of this study may be obscured by the enormous variations in technique that exist even between therapists with very similar training and political affiliation. To understand the process of psychotherapy and psychoanalysis better one would need to study large number of therapists within and between different analytic schools (Kernberg, 1999). Finally, one has to note that the treaters in this study, all analysts themselves, are likely practicing a different sort of psychodynamic psychotherapy than are the non-analysts, who make up the majority of psychoanalytic psychotherapists. It is impossible to generalise the findings here to the larger body of psychoanalytic psychotherapists.

5.5.3 Future research

Future research with the YAWRS to distinguish psychoanalysis from psychodynamic psychotherapy should begin with improvements in the YAWRS measure and sample described in Chapter 4. To increase usability and make it possible to apply it

more widely, the YAWRS should be shortened and careful work needs to be done to test its reliability, with standardisation across therapists and treatments. The validity of the YAWRS should be verified by using it alongside observer and patient-rated measures of process. Next, the YAWRS should be applied to a larger sample of psychoanalytic and psychotherapeutic treatments, beginning with relatively homogeneous clinical perspectives, and expanding to collect data on a wide range of treatments. Larger samples should also be assembled with subgroups by patient pathology, demographics, and other relevant baseline variables. With enough data it will be possible to tease out some of the variables confounded in this study, such as the relationship of treatment length and treatment outcome to process, independent of treatment modality. Lastly, it would be useful to have consistent process data over a longer period of the treatment, so as to separate the process findings of the first year, as discussed above, from those in later phases of psychoanalysis and psychotherapy.

Although the YAWRS has a large number of items covering a wide range of analytically-relevant concepts, there are other variables suggested by the literature which need to be operationalised and measured. Vaughn and Roose (1995) point to the need for a better measure of “working through” and argue that this is an essential component of the analytic process that has not been sufficiently studied. Free association, while measured indirectly in this study, would also benefit from a more careful examination on the YAWRS to test their model. It would be interesting to test their hypothesis that the development of “analytic process” could be traced directly to the existence of four elements: free association, resistance, interpretation, and working through. We suspect that careful measurement of these four attributes would lead to a continuous “analytic process” variable which would be higher in psychoanalysis than in psychotherapy, but with an even rather than a bimodal distribution. Following Kernberg’s model of analytic process, it would also be useful to have more YAWRS items looking at “technical

neutrality” and test how exclusive this is to analytic work. Following Kernberg’s suggestion, a comprehensive, reliable, and well-validated process measure (such as the YAWRS could someday become) could be productively applied to other comparisons within the field, including analytic versus supportive psychotherapy or ego psychological versus object relational vs. self-psychological psychotherapy. Much could be learned from a wider understanding how therapist techniques and patient content differ or remain the same between these modalities.

5.6 Conclusion

In this chapter, the YAWRS, a new analytically-informed therapist-report process measure described in Chapter 4, was applied to a small study comparing the process of psychoanalysis and psychodynamic psychotherapy. A number of hypotheses were proposed based on scant prior empirical findings combined with a large theoretical literature discussing the expected differences between these treatments. A few predictions were also made as to linear trends in process variables. Of the eight differences predicted, five were confirmed using the less-stringent t-test analysis, with trends in the expected direction on the other three. None of the linear trends were observed. We believe that detection of difference between the therapies, despite numerous methodological weaknesses, is evidence in favor of the validity of the YAWRS as a therapist-report measure of outcome and for the usefulness of process measures in studying the differences among treatment modalities. It is unclear whether lack of confirmation of linear trends is due to the insensitivity of the statistical test, the low statistical power in this small study, the lack of observable shifts over a relatively short time period in a heterogeneous set of treatments, or from the fact that treatment process is more uniform over time than is widely described. Weaknesses in the study design include lack of reliability and validity data on the YAWRS, need for other sources of

information in assessing process, small sample size, confounding of the treatment modality distinction with length of treatment and symptomatic improvement, and the fact that all clinicians in this were analysts. Each of these weaknesses calls out for ongoing research, particularly refinement of the YAWRS, validation with other existing measures for which validity data has been collected, and application to larger numbers of subjects in various treatments. Work in this area is methodologically challenging, expensive, and time consuming, but has the promise for greatly elucidating current thinking about analytic work.

CHAPTER 6. THE YOUNG ADULT WEEKLY RATING SCALE: PSYCHOANALYTIC AND PSYCHOTHERAPY OUTCOME

6.1 Introduction

The association of a process measure with therapeutic outcome, done in a careful and meaningful way, is the holy grail of psychotherapy process-outcome research. If the field is able to consistently demonstrate that aspects of the therapeutic process – either on the part of the patient, the therapist, or some interaction of the two – are associated with positive outcome, the stage would be set for empirical confirmation of theorized mechanisms of change, and ultimately to psychotherapy recommendations that are based on objective scientific evidence. In this chapter, we begin by reviewing the process literature, focusing now on data linking process and outcome. Hypotheses based on this literature are outlined. We then apply the YAWRS, as described in Chapter 4 and partially validated in Chapter 5, to the Young Adult sample investigating whether process variables in the first year are associated with positive outcome. As in Chapter 5, the sample size is small, so a number of exploratory hypotheses are entertained and these are discussed in the context of causal explanations and plans for future research.

6.1.1 Theoretical predictions

The theoretical psychoanalytic and psychotherapy literature on mechanisms of change in psychotherapy is large and a full review is beyond the scope of this thesis. In Chapter 2, several theories of structural change and therapeutic process were discussed in the context of which process factors are needed to achieve long-lasting improvement through psychoanalysis or psychotherapy. Several of these themes appeared again in Chapter 5 as differentiating characteristics of psychoanalysis that hopefully make it more suited for bringing about enduring change. For the purposes of this chapter, we will focus on a number of basic attributes of psychotherapy and psychoanalysis that have been theorized, and in most cases, shown to be related to symptomatic improvement. As

the data analysis in this chapter does not involve long-term symptomatic follow-up or a measure of structural change, exclusive emphasis will not be placed on those processes thought to be better at achieving long-term rather than short-term change.

All psychoanalytic theorists of therapeutic action hypothesize that therapy works, at least in part, by helping patients to explore aspects of their emotional and cognitive life of which they have not been fully aware (Fenichel, 1941; Fonagy, 1999b; Greenson, 1994; Jones, 1997; Kernberg, 1999; Modell, 1998). The nature of these explorations differs by clinician and patient but typically includes a number of themes that are found throughout psychoanalytic theories of the mind (Jones & Price, 1998): affect/feelings, unconscious content, defence mechanisms, interpersonal relationships, links between current mental contents and past experience or development, relationship with the therapist, behaviour in the session, dreams, sex, self-image, and guilt. Most psychoanalysis and psychodynamic therapists add that specific elucidation of the way in which the patient recreates other maladaptive relationships in the context of the relationship with the therapist (i.e., the transference) is a core technique by which exploratory processes are facilitated. Verbal interpretation, by the therapist to the patient, of the thoughts and feelings with regard to all these themes, in particular the transference, is usually seen as the central mutative action taken by the clinician. The role of supportive interventions in this process has been hotly debated throughout the history of psychoanalysis, and at times been considered anathema to the essence of psychoanalytic work. However, most recent theorists acknowledge that supportive interventions are necessary at times to maintain the therapeutic alliance, help the patient to feel safe, and allow the more penetrating and difficult work of interpretation to proceed, particularly with personality disordered patients (Fonagy, 1999b; Gabbard et al., 1994; Kernberg, 1999; Wallerstein, 1986).

While most theoretical discussions of psychoanalytic and psychotherapeutic technique have centred around the value of interventions, they implicitly suggest what elements of patient content are signs of productive therapeutic work. If a therapist is successful in helping the patient to explore psychoanalytically relevant content, the themes listed above will appear regularly in the material that the patient presents. Issues of transference and relationships will hold an important place if the therapy is following its course. Most psychoanalytic theorists also acknowledge that an inevitable consequence of deep exploratory work and, in fact, a desired and mutative force in the treatment, is the development of infantile or transference neurosis by the patient, commonly recognized as a regression to more primitive behaviour that can then be studied in the treatment (Fonagy, 1999b; Kernberg, 1999). At the beginning of successful treatments this regression is expected to increase over time. Simultaneously unconscious themes emerge, resistance is discussed and brought under control, and the patient ultimately resolves the neurotic transference in a way that changes basic patterns of thinking and interaction with others.

The question of how the patient and therapist relate to one another during treatment and affectively experience the process is one that is less discussed in the psychoanalytic literature, and on which there is little consensus. Theories abound with a variety of different predictions. Many clinicians, particularly in the psychotherapy literature and in the treatment of patients with poor interpersonal relations, stress how a positive affective experience and good interpersonal alliance experienced both by therapist and patient is necessary for the difficult exploratory work of psychotherapy to proceed (Gabbard et al., 1994; Kohut, 1977; Winnicott, 1962). On the other hand, the material that surfaces in a penetrating treatment is well known to be primitive and emotionally charged (Klein, 1937/1984; Rosenfeld, 1971) and it is expected that this would lead to strong negative affects for the patient (in whom the painful material is

surfacing) and in the therapist (onto whom the affect is usually externalised as part of the transference). According to theory alone, it would appear that the extent and quality of these affective experiences is highly variable by individual characteristics of the patient, the therapist, and the precise stage of treatment.

6.1.2 *Therapist variables*

Fortunately, a large amount of empirical research has been conducted to date on many of the technical and theoretical issues raised above. The Jones PQS (Ablon & Jones, 1998; 1999; Jones & Pulos, 1993), described in Chapters 4 and 5, has succeeded in outlining a set of concrete and codeable therapist interventions that reflect the important components of psychodynamic technique. In three separate studies, Jones and colleagues found that the extent of dynamic technique, as quantified using their measure, was directly related to symptomatic improvement (Ablon & Jones, 1998; 1999; 1993). Most impressively this was true even in groups of patients treated by non-psychoanalytic treatments (two samples of patients in CBT and one in IPT). No other measure to date has comprehensively captured psychoanalytic technique in this way and been applied to both analytic and non-analytic psychotherapies.

A variety of research points to the specific utility of some of the components that Jones folds in to his dynamic technique variable. General measures of therapist interventions have had variable success in capturing these components. The Piper method (ITS) failed to reveal such a relationship in several studies of short term psychotherapy (Hill et al., 1988; Piper et al., 1987). Gaston and Ring, whose measure was developed to focus specifically on the technical ingredients thought to be related to outcome, applied the ITS to three sessions of 16 brief cognitive and dynamic therapies, and found that the “exploration of emotions and cognitions” scale *negatively* predicted outcome (Gaston & Ring, 1992), though was positively associated with alliance in the improved group only. However, in a further study of 120 patients randomized to

behaviour, cognitive, and brief dynamic therapies, exploratory techniques, as identified by the ITS, were found to predict good outcome in all three therapy modalities, in conjunction with poor alliance at mid-therapy in cognitive and behaviour therapy and in conjunction with strong alliance in brief dynamic therapy (Gaston et al., 1998). Bachelor (1991) noted that client-report of exploratory interventions was predictive of positive outcome. In cognitive therapy, discussion of problematic relationships with others has been related to positive outcome (Gaston & Ring, 1992; Orlinsky et al., 1994).

Using the VPPS to collect process and outcome ratings from patient, therapist, and observer, Gomes-Schwartz (1978) and Windholz (1988) showed that exploratory processes and patient involvement correlated with therapist rating of outcome and patient involvement correlated with observer outcome, but no process measures correlated with patient rating of outcome. O'Malley (1983) found that patient involvement also correlated with patient assessment of outcome. Rounsaville and colleagues (1987) showed that VPPS-rated therapist behaviours (therapist exploration and warmth) were significantly predictive of patient outcome, whereas patient behaviours were not.

The positive effect of general therapist interpretations on outcome has been well documented across a wide range of treatments (Castonguay, Goldfried, Wisner, Raue, & Hayes, 1996; Milbrath et al., 1999; Orlinsky, Rønnestad, & Willutzki, 2004; Stiles & Shapiro, 1994). One meta-analysis has suggested that these findings are strongest when patient or observer rate the process, and are inconsistent when rated by therapist (Orlinsky et al., 1994). More in-depth studies show that a complicated relationship exists between therapist interpretations as process and the outcome of treatment. Application of Cooper and Bond's PIRS (see Chapter 4 for description of this measure) to 20 brief psychodynamic psychotherapies showed that patient elaboration of emotional impact was followed by defence interpretation (which in turn was followed by more patient

emotional elaboration), while patient elaboration of the significance of an event was followed by transference interpretation (which was also followed by more patient elaboration of significance). Non-interpretive interventions were followed by patient disclosure of facts, not emotion. Both interpretive intervention sequences and therapist's use of support predicted post-treatment symptom reduction (Milbrath et al., 1999).

A large body of research has been devoted to the psychoanalytic hypothesis that transference interpretations have strong mutative effects in psychotherapy. (Connolly, 1998; Connolly et al., 1999; Gabbard et al., 1994; Garduk & Haggard, 1972; Høglend, 1993c; 1994; 1996; Horwitz et al., 1996; Kernberg et al., 1972; Leary & Gutfreund, 1992; Luborsky, Bachrach, Graff, Pulver, & Christoph, 1979; McCullough et al., 1991; Pessier & Stuart, 2000; Piper et al., 1991; 1993; Silberschatz, Fretter, & Curtis, 1986). While many studies have suggested that transference interpretations can have a negative impact on outcome (Connolly, 1998; 1999; Garduk & Haggard, 1972; Høglend, 1993; Luborsky et al., 1979; McCullough et al., 1991; Silberschatz et al., 1986, Piper, 1991) this has mostly been the case when patient have low levels of object relations, the interpretations come very early in the treatment (Connolly et al., 1999), or when symptomatic outcome is measured immediately after the interpretation (Garduk & Haggard, 1972; Luborsky et al., 1979; McCullough et al., 1991; Silberschatz et al., 1986). Leary and Gutfreund, (1992) showed that interpretations have an initial inhibitory effect but then facilitate progress over the course of an entire session. As part of the Menninger TRIP study (see Chapter 4), Gabbard and colleagues (1994; Horwitz et al., 1996) identified factors that predict a negative response to transference interpretations (neuropsychologically based cognitive dysfunction, history of early trauma, patterns of object relations involving interpersonal distance, masochistic tendencies, and anaclitic rather than introjective psychopathology). In a replication of Kernberg's findings from the Menninger PRP (1972), they found that

transference interpretations were helpful in strengthening the therapeutic alliance and bringing about symptomatic improvement, if used after supportive interventions.

Supportive interventions have also been studied as independent predictors of improvement in psychodynamic psychotherapy. Using patient-rated measures of process, derived from Orlinsky and Howard's classic Therapy Session Report (Orlinsky & Howard, 1967; 1986), Kolden demonstrated that therapeutic bond and therapeutic realizations (which included such supportive interventions as unburdening, encouragement, and mastery) predict positive outcome when rated in early sessions across different treatment modalities and patient presentations (Kolden, 1991; 1996a; 1996b; 1992; 1996). Bachelor (1991) reported that client report of therapist "helpfulness" was predictive of patient problems, and meta-analyses have also suggested that therapist attention and support has beneficial effects (Orlinsky et al., 1994; Orlinsky et al., 2004). Other studies, cited above, have demonstrated the important role of supportive work in the context of creating an alliance for more penetrating interventions (Gabbard et al., 1994; Gaston, Piper, Debbane, Bienvenu, & Garant, 1994; Gaston et al., 1998; Kernberg et al., 1972; Rounsaville et al., 1987).

One of the most difficult aspects of therapist behaviour to study is the nature of the therapist's affective response to the patient in the context of the treatment. As outlined above, theoretical work has not made a specific prediction about the relationship between therapist affect and outcome. Empirical research is similarly inconsistent in its findings. Rounsaville (1987) found that therapists scored higher on the VPPS scale for "therapist warmth" had higher rates of success, but therapist "negative attitude" showed no relationship. More specific ratings of therapist affective response, such as the use of "emotion words," particularly those related to anger, have been found to predict positive outcome as well (Hölzer et al., 1997). Given this range, it seems likely

that a measure of therapist positive view of treatment, will not on its own be predictive of outcome.

6.1.3 Patient variables

Measures of patient content and behaviour in therapy have been less successful in predicting outcome, though many of the theories suggesting therapist findings have associated predictions with patients (Milbrath et al., 1999). While created as a measure to characterize therapist interventions, the Jones PQS may be easily adapted into a rating of patient content (Ablon & Jones, 1998; Jones & Pulos, 1993). If the therapist is to comment on feelings, unconscious content, defences, links to the past, transference, in-session behaviour, dreams, sex, self-image, and guilt, the patient must be producing material relevant to these themes. The YAWRS, as described in Chapter 4, is uniquely suited by its structure to capture the extent of patient discussion for the very same items on which the analyst intervention was rated. We would expect there to be association between the two ratings, but that each would provide a distinct contribution to the prediction of treatment success.

Similarly, if we predict that transference interpretations are mutative and lead to symptomatic change, it is also likely that the transference themes appear more in successful treatments. In earlier decades of psychoanalytic research, much effort was devoted to the question of whether resolution of the transference neurosis is necessary for a successful treatment. To the surprise of investigators in several studies, evidence of ongoing transference material was seen even in otherwise successful cases (Bachrach, 1993; Erle, 1979; Norman et al., 1976; Oremland et al., 1975; Schlessinger & Robbins, 1974). This was early evidence that transference continues throughout successful cases and may be evidence of a good treatment. When Graff and Lubosky's analyst-completed checklist was applied to two successful analyses, the measure showed a trend towards higher transference and lower resistance. In two less successful treatments, transference

and resistance were closely associated and a drop in resistance was not noted (Graff & Luborsky, 1977).

An increase in the extent of patient regression in successful psychoanalytic treatments is a theoretically predicted process result which has never been directly tested or directly measured. A number of psychoanalytic studies have noted, somewhat to their surprise, a finding that indirectly may point to this pattern. Wallerstein (1986), Erle (1979), Kantrowitz, (1993), and Sandell (2000) all note the identification, and in some cases initial appearance, of increasing psychopathology during the early course of successful treatments. Further information has not been available about the fate of these individual symptoms and how they were related to the material raised by the patient. In a time-series analysis of psychoanalytic treatment in a brittle diabetic patient, Moran and Fonagy (1991) showed that relevant interventions can directly lead to symptom exacerbation, before structural change occurs, eventually followed by symptom improvement. With the frequent collection of YAWRS data it should be possible to show a trend towards increasing regression in the first year as a predictor of treatment success.

An ideal measure of patient material to identify successful treatments, from a theoretical perspective, would be one that captures both the quality of material coming from the patient and its relationship to the level of resistance. By all the theories of psychoanalytic process reviewed above (Fenichel, 1941; Fonagy, 1999b; Greenson, 1994; Jones, 1997; Kernberg, 1999; Modell, 1998) a higher ratio of useful patient material to resistance should be predictive of better treatment outcome. Wallner Samstag and colleagues (Samstag, Batchelder, Muran, Safran, & Winston, 1998) reported a similar result, using the Session Evaluation Questionnaire (SEQ Stiles, 1980). The SEQ asks therapist and patient to choose between 22 bipolar adjective pairs in describing a session. Factor analysis of ratings showed two independent factors for each: depth/value and

smoothness/ease. Wallner Samstag found that treatment adherence and success were associated only with patient rating of depth and therapist rating of smoothness (Samstag et al., 1998). Interestingly, the global factor analysis of YAWRS subsection summary scores described in Chapter 4 produced patient factors for clear unconscious themes and resistance, roughly analogous to the two scales of the SEQ. We predict that the ratio of these two factors during the first year will be the best predictor of patient improvement and that the factor itself will increase over time more for improvers, reflecting successful analysis of resistance.

Patient's positive view of treatment appears to be somewhat more useful in judging the success of treatment than the therapist's view. Several of the positive factors from a patient's representation of therapy and therapist (from the Intersession Experience Questionnaire and the Therapist Involvement Scale, both designed by Orlinsky) correlate with patient estimates of therapeutic benefit (Orlinsky & Geller, 1993). In a meta-analysis, Orlinsky reports studies with a range of modalities pointing to the predictive value of patient view of therapy (Orlinsky et al., 1994; 2004).

6.1.4 Therapist-patient interaction variables

Measures of patient-therapist alliance have historically been the strongest and most consistent predictors of positive therapeutic outcome. Results have varied slightly with different measures. Numerous published studies using Horvath and Greenberg's WAI, including a meta-analysis, found excellent internal consistency of the measure and a "moderate but reliable" association between high ratings of working alliance and positive treatment outcomes across different types of therapy and rating sources. These results were strongest for client assessment of alliance and weakest for observer reports (Horvath & Symonds, 1991). Results of empirical studies looking at the factor structure of WAI alliance ratings have varied, with some concluding that a "general alliance" factor is paramount, leaving goal, task, and bond as subfactors (Tracey & Kokotovic, 1989), and

others suggesting that measure of the relationship (i.e., bond) can be definitively separated from a measure of client agreement and confidence (associated with goal and task), particularly in CBT (Andrusyna et al., 2001; Hatcher & Barends, 1996).

Luborsky's HAQ-II has been found to be closely associated with other measures of alliance (Luborsky et al., 1996) and with quality of object relations and treatment outcome (Piper et al., 1991). Gunderson's prospective study of treatment outcome in eclectic psychotherapy of 33 female patients with borderline personality disorder used the HAQ to reach six conclusions: (1) alliance was rated highly by patients and therapists throughout treatment, (2) alliance showed steady and significant improvement over time, (3) patients and therapists corresponded closely in their assessments of alliance, (4) therapists rated the alliance higher than did patients at three and four years, (5) therapist ratings of alliance at six weeks were predictive of subsequent dropout, but (6) early alliance scores were not strongly related to subsequent level of change (Gunderson, Najavits, Leonhard, Sullivan, & Sabo, 1997).

Two studies have looked specifically at the role of therapeutic alliance as measured by the CALPAS in the treatment of personality disorders. Bond and colleagues (1998) applied the CALPAS and PIRS (described above as a measure of therapist intervention) to transcripts of five subjects with borderline or narcissistic personality disorder and found that transference interpretations led to a deterioration of alliance when the alliance was weak, but to enhanced work when the alliance was solid. Marziali and colleagues (1999) randomly assigned 110 borderline patients to interpersonal group psychotherapy and individual dynamic psychotherapy, measuring symptoms, social adjustment, service utilization, and alliance at regular intervals. In the individual psychotherapy cohort, early and late alliance ratings were related to outcome at 12-month follow-up but not at 24 months. In the group psychotherapy cohort, alliance was unrelated to outcome at 12 month follow-up but was significantly associated with outcome at 24 months. The

authors conclude that alliance is a relevant measure for predicting outcome in the borderline population but the subtleties of the relationships remain to be adequately explored. They speculate that the failure of association with 24-month followup in individual therapy is due to inadequate power and that the failure of association with 12-month followup in group therapy shows that treatment effects may take longer to consolidate in this modality (Marziali et al., 1999).

A few studies have looked at the correspondence between various measures of therapeutic alliance and at their differential success in predicting treatment outcome. Tichenor and Hill (1989) compared observer rated versions of the CALPAS, HAQ, VTAS, and WAI (WAI-O), as well as client and therapist versions of the WAI (WAI-C and WAI-T, respectively), in a sample of eight depressed outpatients undergoing brief psychotherapy. All measures demonstrated high internal consistency and interrater reliability. Strong associations between CALPAS, VTAS, and WAI-O were found, with some correspondence between WAI-O and HAQ. WAI-C and WAI-T were not significantly related to each other or to any of the observer-rated measures. Bachelor (1991) applied patient and client versions of the HAQ, TARS, and VPPS to the eclectic but largely humanistic psychotherapy of 47 female clients of a university counseling service. She reported that patient and therapist agreed on ratings of positive toned therapist contributions (except for therapist exploration) but differed on their perceptions of most other alliance variables. Client perceptions of therapist-offered helpfulness, warmth, emotional involvement, and exploratory interventions were the best predictors of positive treatment outcome. Safran and Wallner (1991) applied patient versions of the WAI and CALPAS to a sample of 22 patients in short-term cognitive therapy and found that scores were equally predictive of global outcome, with the CALPAS slightly better at predicting change in individual outcome measures. Fenton and colleagues (Fenton, Cecero, Nich, Frankforter, & Carroll, 2001) used observer versions

of the CALPAS, HAQ, VTAS, and WAI, as well as client and therapist versions of the WAI, to compare predictive validity in a sample of 46 patients undergoing cognitive or twelve-step manualized therapy for substance use. The four observer rated scales predicted outcome equally well, whereas the WAI-C and WAI-T were not associated with client improvement.

Ackerman and Hilsenroth (2003) review the existing literature on the relationship of therapist characteristics and technique to therapeutic alliance. They found that therapist's personal attributes such as being flexible, honest, respectful, trustworthy, confident, warm, interested and open have all shown significant effects on the alliance. Meanwhile, therapist techniques including exploration, reflection, noting past therapy success, accurate interpretation, facilitating the expression of affect, and attending to the patient's experience also contribute positively to alliance.

Finally, Orlinsky's meta-analyses (1994; 2004), which include studies relating psychotherapeutic process to outcome across a wide range of treatment modalities, have noted a relationship between contract discussion and treatment success. While most obviously applicable to short-term cognitive behavioural treatments in which the discussion of the terms of treatment is emphasized, it would be interesting to test the value of this approach to psychoanalytic treatments as well.

6.2 Hypotheses

Process variables measured in the first year of treatment will successfully predict which patients will go on to be symptomatic improvers versus non-improvers at the end of treatment.

Therapist variables

1. Extent of dynamic technique as operationalised by Jones (1993; Ablon & Jones, 1998) will be greater in improvers.

2. Overall extent of interpretation will be greater in improvers (Jones, Ghannam et al., 1993; Jones & Price, 1998; Milbrath et al., 1999; Orlinsky et al., 1994).
3. Extent of transference interpretations will be greater in improvers (Kernberg et al., 1972).
4. Extent of interpretations related to the patient's external relationships will be greater in improvers (Gaston & Ring, 1992).
5. Extent of exploratory interventions will be greater in improvers (Bachelor, 1991; Gomes-Schwartz, 1978; Rounsaville et al., 1987; Windholz & Silberschatz, 1988).
6. Extent of supportive interventions will be greater in improvers (Bachelor, 1991; Kernberg et al., 1972; Kolden, 1991; 1996a; 1996b; 1992; 1996; Rounsaville et al., 1987).
7. Therapist's countertransference will *not* be related to treatment outcome (Hölzer et al., 1997; Rounsaville et al., 1987).

Patient variables

8. Extent of dynamic material, as operationalised by Jones (1993; Ablon & Jones, 1998) will be greater in improvers.
9. Extent of transference themes will be greater in improvers (Bachrach, 1993; Erle, 1979; Graff & Luborsky, 1977; Norman et al., 1976; Oremland et al., 1975; Schlessinger & Robbins, 1974) and increase over time (Graff & Luborsky, 1977).
10. Extent of relationship material will be greater in improvers (Orlinsky et al., 1994).
11. Extent of regression will increase over time in improvers (Erle, 1979; Kantrowitz, 1993; Sandell et al., 2000; Wallerstein, 1986).
12. Ratio of clear unconscious themes to resistance will be higher in improvers and increase over time in improvers (Samstag et al., 1998).
13. Patient's positive view of treatment will be higher in improvers (Orlinsky & Geller, 1993; Samstag et al., 1998)

Patient-therapist interaction

14. Patient-therapist contract will be discussed more with improvers (Orlinsky et al., 1994).
15. Patient-therapist alliance will be higher with improvers (Horvath & Symonds, 1991; Horwitz, 1974; 1996a; Kolden, 1996b; 1996; Marmar, Gaston et al., 1989; 1989; Marziali et al., 1999; Orlinsky et al., 1994; Rudolf, 1991).

6.3 Methods

The basic experimental design, including subject and assessments, has been described in Chapter 3, 4, and 5. Data analysis was performed on the 10 Young Adult subjects for whom at least 6 months of YAWRS data was collected during the first year of analytic treatment. As was demonstrated in Chapter 3, there were no significant differences found in demographic or diagnostic variables between the 10 subjects in this analysis and the 14 subjects that were excluded.

6.3.1 Data processing

The final step in preparation of the YAWRS data for analysis entailed accommodating the empirically derived factors and scales derived to this point with the theoretical intent of the hypotheses presented in the introduction. Each theoretical variable mentioned in one of the hypotheses was matched with the derived global factors, YAWRS factors, subsection summary scores, and individual YAWRS items, in that order of preference. The minimum number of factor analytic, subsection summary, and item scores with the greatest breadth were selected to correspond to each hypothesis scale, so as to cover the concept adequately and use empirically derived means for maximizing scale reliability and coherence.

6.3.2 *Data analysis*

The relationships of process scales to treatment intensity and outcome were tested using two data analytic methods. In the first, more descriptive method, mean values were calculated on each of the 15 scales for each subject on all available YAWRS data collected during the first year of treatment. The number of values that entered into these means varied between 6 (for two psychotherapy subjects for whom data from only half a year was available) and 42 (for a psychoanalytic subject who had attended regularly all year). The mean number of YAWRS data points per subject was 20.1 with a standard deviation of 15.5. Student's independent samples t-tests were performed to test the difference between psychoanalytic and psychotherapy cases for the scales suggested in Hypothesis 3 and between symptomatic improvers and non-improvers for the scales suggested in Hypothesis 4. Due to the exploratory nature of this analysis and the fact that individual predictions were made for each of these analyses, Bonferroni corrections were not used and the α was set at 0.05.

In the second method, an attempt was made to compare linear trends as well as mean differences between psychotherapy and psychoanalytic patients and symptomatic improvers and non-improvers on the 15 scales. Hierarchical linear modeling (HLM), a technique developed for comparing irregular data sets in a nested structure, such as time series data for a series of subjects divided into comparison groups, was selected for this purpose (Bryk & Raudenbush, 1987; 2001; 1996). The HLM/2L software procedure uses an iterative technique to model the relationship between group membership (intensity of treatment or symptomatic improvement) and linear parameters of YAWRS scales taken over time (slope and intercept). Because the psychoanalytic and improvement groups only differ by a single subject and the HLM procedure is optimized for finding differences between groups (which in this case would be highly subject to the characteristics of that single subject), the HLM procedure was applied separately for the

two grouping variables to each scale. A fixed effect estimate of the relationship between group membership and intercept or slope that was significantly different from 0 in the direction predicted was considered to be a significant finding. As above, because of the exploratory nature of these analyses and the specificity of the predictions, Bonferroni corrections were not used and the α was set at 0.05.

6.4 Results

6.4.1 Generation of hypothesis-based scales

Based on the results of the subsection and subsection summary factor analyses, sets of variables were chosen to correspond to the process constructs outlined in the introduction and hypotheses of this chapter. In order to maximise the empirical basis for these scales, yet cover the constructs adequately, scales were assembled with the following order of preference: (1) global factors (GLF), (2) subsection factors (FAC), (3) subsection summary scores (SSC), and (4) individual YAWRS items (ITM). The composition of the seven therapist scales is outlined in Table 5.3. The six patient scales are described in Table 5.4. Formulae for all 15 scales are given in Appendix 5.2d.

1. Jones dynamic technique

- a. Interprets feelings: interprets sadness [PC2oi, SSC], primitive emotional stance [PC3oi, SSC], and anger [PC3oi, SSC]
- b. Interprets unconscious: interprets general unconscious [UAoi, SSC], unconscious reactions to aggression [UBoi, SSC], unconscious content related to sexuality [UCoi, SSC], unconscious content related to self and self-esteem [UDoi, SSC], and unconscious content related to body [UEoi, SSC]
- c. Interprets defences: analyst comments on defensive manoeuvres [UG15, ITM], patient's resistance [UG16, ITM], patient's relationship to analytic process [UG17, ITM], interprets defences: projection, reaction formation, denial [PDF1i, FAC], isolation [PDF2i, FAC], splitting [PDF3i, FAC], regression, projective identification [PDF4i, FAC], and externalisation [PDF5i, FAC].
- d. Interprets links to past: interprets historical material [MCOi, SSC], comments about patient's experiences of little child or baby within [UG12, ITM]
- e. Interprets transference: see item 3 below.
- f. Interprets in-session behaviour: interprets resistance [GAoi, SSC], general attitude to analysis [GBoi, SSC], time keeping [GCoi, SSC], missed sessions [GDoi, SSC], patient's behaviour in session [GEoi, SSC], patient's aggression and sexuality in the analysis [GGoi, SSC].
- g. Interprets clear themes: interprets clear theme to material [GF02i, ITM], completeness and quality of analyst's understanding of patient's material [J3, ITM], analyst comments on links between behaviour outside the session to current treatment material [UG13, ITM].
- h. Interprets dreams: interprets dreams [GF05i, ITM], daydreams and fantasies [GF06i, ITM], associations to dreams and fantasies [GF07i, ITM], meaning or significance of thoughts and fantasies [GF08i, ITM].
- i. Interprets sex: interprets sexual relations [MFOi, SSC], sexuality [MKoi, SSC], unconscious themes related to sexuality [UCoi, SSC].
- j. Interprets self-image: interprets manifest themes related to body [MAoi, SSC] and self-esteem [MBoi, SSC], interprets unconscious themes related to self and self-esteem [UDoi, SSC] and unconscious content related to body [UEoi, SSC].
- k. Interprets guilt: interprets manifest guilt, punishment, and conflicted loyalty themes related to historical material [MC04ci, ITM], family [MD04bi, MDo4hi, ITM], friends [ME04ci, ME04ji, ITM], sexual relations [MF04di, MF04ji, ITM], sexuality [MK06di, ITM], interprets guilt concerning sexual wishes [PC02ci, ITM], aggressive wishes and actions [PC02di, ITM], striking absence of guilt [PC03fi, ITM], fear of being humiliated [PC03ii, ITM], intense shame [PC03ji, ITM], and overwhelming guilt in relation to analyst [UF04bi, ITM].

Table 6.3. Composition of therapist-based hypothesis scales (continued).

-
2. General interpretation: all interpretation of general, manifest, preconscious, and unconscious content [GAoi-GHoi, MAoi-MLoi, PA1oi-PDf5i, UAoi-UF4i, UGf1-UGf4, FAC and SSC].
 3. Transference interpretation: interprets general transference [PA1oi, SSC], positive wishes towards analyst [PA2oi, SSC], transference with anxiety [PA3oi, SSC], transference with competitive and aggressive themes [PA4oi, SSC], transference with resentment [PA5oi, SSC], and primitive transference [PA6oi, SSC], interprets change in transference across week [PBoi, SSC], analyst comments on connections between therapeutic relationship and past relationships [UG09, ITM], connections between therapeutic relationship and present relationships [UG10, ITM], and displacing feelings from the analyst to outside figure [UG11, ITM].
 4. Relationship interpretation: interprets relationship themes within family [MDoi, SSC], with friends [MEoi, SSC], sexual relations [MFoi, SSC], sexuality [MKoi, SSC], links between behaviour outside session to current treatment material [UG13, ITM], and possible meaning of others' behaviour [UG13, ITM].
 5. Exploratory interventions: interprets thinking [UGf2, FAC]
 6. Supportive interventions: supportive intervention [UGf3, FAC]
 7. Therapist's positive view of treatment: analyst feels patient was helped [UF3p, SSC], analyst feels loving/empathic [UHf3o, FAC], analyst's judgement of "quality of week" [Javg, SSC], - analyst feels attacked/disgusted [UHf1o, FAC], - inadequate/confused [UHf2o, FAC], and - bored/cutoff [UHf4o, FAC], - patient is bullying [UIf1o, FAC], - narcissistic [UIf2o, FAC], and - analyst is rejected [UIf3o, FAC].
-

Table 6.3. Composition of therapist-based hypothesis scales.

-
1. Jones dynamic material
 - a. Feelings: sadness [PC2op, SSC], primitive emotional stance [PC3op, SSC], anger [PC3op, SSC]
 - b. Unconscious content: general [UAop, SSC], reactions to aggression [UBop, SSC], related to sexuality [UCop, SSC], related to self and self-esteem [UDop, SSC], related to body [UEop, SSC]
 - c. Defences: projection, reaction formation, denial [PDf1p, FAC], isolation [PDf2p, FAC], splitting [PDf3p, FAC], regression, projective identification [PDf4p, FAC], externalisation [PDf5p, FAC].
 - d. Links to past: historical material [MCOp, SSC]
 - e. Transference: see item 2 below.
 - f. In-session behaviour: resistance [GAop, SSC], general attitude to analysis [GBop, SSC], time keeping [GCop, SSC], missed sessions [GDop, SSC], bad behaviour in session [GEop, SSC], patient's aggression and sexuality in the analysis [GGop, SSC].
-

Table 6.4. Composition of patient-based hypothesis scales (continued).

-
- g. Clear themes: clear theme to material [GF02p, ITM].
 - h. Dreams: dreams [GF05p, ITM], daydreams and fantasies [GF06p, ITM], associations to dreams and fantasies [GF07p, ITM], meaning or significance of thoughts and fantasies [GF08p, ITM].
 - i. Sex: sexual relations [MFop, SSC], sexuality [MKop, SSC], unconscious themes related to sexuality [UCop, SSC].
 - j. Self-image: manifest themes related to body [MAop, SSC] and self-esteem [MBop, SSC], unconscious themes related to self and self-esteem [UDop, SSC] and unconscious content related to body [UEop, SSC].
 - k. Guilt: manifest guilt, punishment, and conflicted loyalty themes related to historical material [MC04cp, ITM], family [MD04bp, MDo4hp, ITM], friends [ME04cp, ME04jp, ITM], sexual relations [MF04dp, MF04jp, ITM], sexuality [MK06dp, ITM], interprets guilt concerning sexual wishes [PC02cp, ITM], aggressive wishes and actions [PC02dp, ITM], striking absence of guilt [PC03fp, ITM], fear of being humiliated [PC03ip, ITM], intense shame [PCO3jp, ITM], and overwhelming guilt in relation to analyst [UF04bp, ITM].
2. Transference themes: general [PA1op, SSC], positive wishes towards analyst [PA2op, SSC], transference with anxiety [PA3op, SSC], transference with competitive and aggressive themes [PA4op, SSC], transference with resentment [PA5op, SSC], primitive transference [PA6op, SSC], change in transference across week [PBop, SSC].
 3. Relationship themes: within family [MDop, SSC], with friends [MEop, SSC], sexual relations [MFop, SSC], and sexuality [MKop, SSC].
 4. Regression: aggression and sexuality in the analysis [GGop, SSC], immaturity of patient's mental functioning [GHop, SSC].
 5. Ratio of unconscious themes to resistance: [SOFNrat, GLF]
 6. Patient's positive view of treatment: positive wishes towards analyst [PA2op, SSC], patient feels helped [UF1p, SSC], - negative attitude [GBop, SSC], - transference with anxiety [PA3op, SSC], - transference with competitive and aggressive themes [PA4op, SSC], - transference with resentment [PA5op, SSC], - primitive transference [PA6op, SSC], - patient reports not feeling helped [UF2p, SSC], - analyst feels patient reacts negatively [UF4p, SSC].
-

Table 6.4. Composition of patient-based hypothesis scales.

-
1. Discussion of patient-therapist contract: discussion of treatment parameters [MLop, SSC], interpretation of treatment parameters [MLoi, SSC]
 2. Patient-therapist alliance: patient's general stance (therapeutic alliance) [J1, ITM]
-

Table 6.5. Composition of interaction-based hypothesis scales.

6.4.2 Association between YAWRS scales and symptomatic improvement

Table 6.8 shows the results of the mean comparisons based on the 15 hypotheses. Of the 13 predicted mean differences (excluding therapist's positive view of patient and regression), eight were found to be statistically significant in the expected direction. Jones dynamic technique, general interpretation, transference interpretation, Jones dynamic material, transference themes, relationship themes, ratio of clear unconscious themes to resistance, and contract discussion were all significantly higher in symptomatic improvers, as recorded by the analyst during the first year of treatment. Four of the remaining five variables fell along the predicted pattern (though did not reach the level even of a trend), while only patient's positive view of treatment showed a pattern opposite from predicted (but also did not reach the level of a trend).

	Mean (SE)		Student's T-test (df=8)
	Improver (n=4)	Non-improver (n=6)	
Therapist variables			
(a) Jones dynamic technique	0.89 (0.11)	0.59 (0.06)	2.61*
(b) General interpretation	0.73 (0.11)	0.40 (0.08)	2.49*
(c) Transference interpretation	0.94 (0.07)	0.51 (0.08)	3.72**
(d) Relationship interpretation	0.80 (0.11)	0.39 (0.09)	2.82*
(e) Exploratory interventions	1.58 (0.31)	1.30 (0.25)	0.72
(f) Supportive interventions	0.70 (0.26)	0.29 (0.16)	1.42
(g) Therapist's positive view of treatment	1.95 (0.21)	2.18 (0.09)	-1.09
Patient variables			
(h) Jones dynamic material	0.92 (0.08)	0.58 (0.05)	3.56**
(i) Transference themes	0.76 (0.07)	0.33 (0.08)	3.75**
(j) Relationship themes	0.66 (0.08)	0.28 (0.04)	4.7**
(k) Regression	0.76 (0.16)	0.41 (0.16)	1.48
(l) Ratio of clear unconscious themes to resistance	0.94 (0.09)	0.61 (0.04)	3.77**
(m) Patient's positive view of treatment	1.99 (0.04)	2.19 (0.10)	-1.61
Patient-therapist interaction			
(n) Contract discussion	0.65 (0.09)	0.30 (0.07)	3.23*
(o) Alliance	1.60 (0.12)	1.38 (0.15)	1.05

Table 6.8. Results of T-test analysis of Hypothesis 4: Improvers versus non-improvers.

† p < 0.1, * p < 0.05, ** p < 0.01

Table 6.9 contains the results of the HLM analysis of Hypothesis 4. A significant effect of treatment intensity on the intercept of variables plotted against time over the first year was found for transference interpretation, Jones dynamic material, and transference themes, all in the expected direction. Trends in the predicted direction were found for general interpretation, relationship themes, and ratio of clear unconscious themes to resistance. Of the three predicted interactions between symptomatic improvement and change over time (Jones dynamic technique, general interpretations,

and regression), none were found to be significant. Two unpredicted findings with regard to slope were detected: Jones dynamic material decreased over time in both improvers and non-improvers and supportive interventions increased over time more in the improvers than in the non-improvers.

	Fixed effect estimate (SE)		
	Symptomatic improvement → intercept	Baseline slope	Symptomatic improvement → slope
Therapist variables			
(a) Jones dynamic technique	0.27 (0.15)	0.000 (0.002)	0.001 (0.002)
(b) General interpretation	0.31 (0.17) [†]	0.000 (0.001)	0.001 (0.002)
(c) Transference interpretation	0.56 (0.19)*	0.002 (0.003)	-0.005 (0.004)
(d) Relationship interpretations	0.44 (0.16)*	0.004 (0.003)	-0.006 (0.004)
(e) Exploratory interventions	0.48 (0.54)	0.003 (0.006)	-0.007 (0.009)
(f) Supportive interventions	0.15 (0.25)	-0.002 (0.003)	0.009 (0.004)*
(g) Therapist's positive view of treatment	-0.33 (0.26)	-0.001 (0.004)	0.004 (0.007)
Patient variables			
(h) Jones dynamic material	0.28 (0.12)*	-0.004 (0.001)*	0.002 (0.002)
(i) Transference themes	0.48 (0.17)*	0.002 (0.002)	-0.002 (0.004)
(j) Relationship themes	0.29 (0.14) [†]	-0.004 (0.003)	0.003 (0.004)
(k) Regression	0.39 (0.28)	0.002 (0.003)	-0.002 (0.004)
(l) Ratio of clear unconscious themes to resistance	0.26 (0.13) [†]	-0.001 (0.002)	0.002 (0.004)
(m) Patient's positive view of treatment	-0.32 (0.19)	-0.002 (0.003)	0.004 (0.005)
Patient-therapist interaction			
(n) Contract discussion	0.32 (0.15) [†]	0.002 (0.002)	0.000 (0.003)
(o) Alliance	-0.13 (0.41)	-0.007 (0.008)	0.013 (0.012)

Table 6.9. Results of HLM analysis of Hypothesis 4: Improvers versus non-improvers.

[†] $p < 0.1$, * $p < 0.05$, ** $p < 0.01$

6.5 Discussion

6.5.1 *Therapist variables*

The findings above broadly support the hypothesis that therapist-reported process from the first year of treatment can be useful in predicting symptomatic outcome in psychodynamic psychotherapy and psychoanalysis. By the less stringent t-test methodology, the therapist dynamic technique variable, as derived from the Jones PQS and based on factor analyses described in Chapter 4, predicts positive outcome. This supports the work of Jones and colleagues (Ablon & Jones, 1998; 1999; 1993) linking observer-rated measurement of analytic-process to outcome in psychotherapy, broadening it to include therapist ratings and psychoanalysis. The result is a demonstration, in contrast with the claims of Piper and Hill (Hill et al., 1988; 1987) that general measures of technique can usefully differentiate improvers from non-improvers. The finding also non-specifically supports the theoretical claim that dynamic technique is central to therapeutic action and mechanism for patient improvement (Fenichel, 1941; Fonagy, 1999b; Greenson, 1994; Jones, 1997; Kernberg, 1999; Modell, 1998). It is limited, however, by the large overlap in this study between those subjects in psychoanalysis (who were demonstrated in Chapter 5 to have had more dynamic technique) and those who improved, and by the small sample size studied.

Interestingly, the more specific measurement of “exploratory interventions” on the YAWRS was not associated with outcome, just as it did not show a difference in Chapter 5 between psychoanalysis and psychodynamic psychotherapy. This contrasts with the findings from observer-rated “exploratory technique” using the VPPS (Gomes-Schwartz, 1978; O'Malley et al., 1983; Windholz & Silberschatz, 1988), ITS (Gaston et al., 1998), and PIRS (Milbrath et al., 1999). Several factors likely contribute to this discrepancy. First, the “exploratory interventions” measure used in this study was drawn from a single factor, based on three YAWRS items, limiting the reliability and validity that can come

from a larger number of items. Second, the range in “exploratory” ratings in this sample was small, probably because all therapists were using such techniques, and this measure was unable to capture the detailed differences in degree of analytic technique as was done by the “analytic technique” scale. Finally, while clinicians may reliably and validly report their practices when defined specifically, they are probably not good at generalizing from these practices to say how “exploratory” they have actually been (as the items in this scale request). Other studies that have shown the importance of “exploratory technique” have done so using observer-ratings.

A significant finding in the T-test analysis and a near-significant trend using HLM supports the prediction that a general measure of therapist interpretation is predictive of good therapeutic outcome in this sample, as reported by several prior studies and reviews (Castonguay et al., 1996; Jones, Ghannam et al., 1993; 1998; Milbrath et al., 1999; Orlinsky et al., 1994; Stiles & Shapiro, 1994). This also agrees with the universal psychoanalytic theoretical argument that interpretation in the context of the therapeutic relationship is essential to change (Fonagy, 1999b; Jones, 1997; Kernberg, 1999). The original aspect of this finding is that it argues against prior researchers who claim that therapist ratings of interpretation are biased and less useful in predicting outcome (Orlinsky et al., 1994). The YAWRS general interpretation measure collects information in a broad way on whether analysts interpreted specific themes that they had rated as present, therefore reducing the subjectivity of the response. Also the focus on dynamic interpretations by a sophisticated rater is likely more useful in predicting outcome than less well-trained therapists rating non-dynamic interpretations.

The findings of increased transference and relationship interpretations in the improver group is the strongest YAWRS-outcome result and a significant confirmation of prior theoretical (Fonagy, 1999b; Kernberg, 1999) and empirical work (Gabbard et al., 1994; Horwitz et al., 1996; Kernberg et al., 1972; Leary & Gutfreund, 1992). This

contrasts with the findings of many studies that transference interpretations have a negative impact on outcome (Connolly, 1998; 1999; Garduk & Haggard, 1972; Høglend, 1993; Luborsky et al., 1979; McCullough et al., 1991; Piper, 1991; Silberschatz et al., 1986). A number of differences between the sample and methodology of this study, as opposed to studies that showed a negative effect, are likely responsible. First, outcome in those studies was frequently measured in session or shortly after the interpretation, whereas in this study outcome was measured at termination which was up to 7 years later than the process measure was collected (Leary & Gutfreund, 1992). Second, the analytic clinicians in the present study were more likely than some of the less well-trained therapists in previous studies to keep transference interpretations appropriate to the alliance and level of object relations in the patient (Cabbard et al., 1994). Finally, it is likely that transference interpretations are more likely to be tolerated and useful in the context of open-ended long-term treatments, as opposed to time-limited therapies, as have often been studied (Connolly, 1998; 1999; Luborsky et al., 1979; McCullough et al., 1991).

Interestingly, the prediction of outcome based on level of transference interpretation, as reported by the therapist, was successful despite two potential drawbacks of this study: the fact that transference interpretations were measured only in the first year of treatment (a time when they would potentially be premature and threatening to the alliance) and that no consideration was given (due to small sample size) to different interaction patterns or levels of patient pathology. This suggests that the importance of transference interpretation is a robust finding that would only be enhanced by more sophisticated analyses. It is also worth considering that a therapist's report of transference interpretation may provide early data on the suitability of a patient for psychoanalytic treatment, engagement with the therapist, and clarity of their key dynamic issues. All three contribute to the likelihood of a therapist making transference

interpretations in the first year and may independently be associated with positive outcome. Therapists were less likely to report transference interpretations while treating patients with more primitive psychopathology or lack of engagement, and these patients were less likely to improve because of what they brought to the treatment, independently of the therapist intervention. On the other hand, transference interpretations were significantly higher in the psychoanalytic group as compared to the psychotherapy group (see Chapter 5), though the two groups did not differ significantly in pathology, reducing the likelihood that the association between interpretations and outcome is completely due to confounders.

There is significant evidence that measurement of interaction sequences involving transference interpretations is necessary for understanding the complicated link between these interventions and outcome (Gabbard et al., 1994; Høglend, 1993c; Horwitz et al., 1996; Milbrath et al., 1999; Piper et al., 1991; 1993). The homogeneity of technique and patient pathology in the current sample may explain why such measurement was not necessary for this finding. Future research with larger samples is necessary to further explore this link.

While the lack of association between supportive interventions and outcome differs from the prediction based on basic theory and findings from studies of psychotherapy process, it is readily explainable given some of the results presented in Chapter 5 and earlier in this section. As discussed in Chapter 5, the extent of supportive interventions did not distinguish between psychodynamic psychotherapy and psychoanalysis because these interventions are found in psychotherapies of all theoretical orientations due to their role in facilitating alliance and providing a framework for more penetrating interventions (Ablon & Jones, 1998; Jones & Pulos, 1993; Kernberg, 1999). In the present study's small and homogeneous sample and treatment, the range of supportive interventions was not able to capture the subtle interaction by which support facilitates a

working alliance and allows transference interpretations to be effective. Better measurement of interaction techniques, might make it possible to better note the differential contribution of supportive interventions (Gabbard et al., 1994; Gaston et al., 1994; Gaston et al., 1998; Kernberg et al., 1972; Rounsaville et al., 1987).

As with the measurement of “exploratory interventions” the supportive intervention scale was drawn from a single factor based on only four YAWRS items which directly query the analyst about their interventions. Clinicians are likely poor at noting small differences when asked generally about their technique, and the validity and reliability of this scale is in question. Finally, if we consider the confounding variables that may account for some or all of the association between interpretation and outcome (i.e., more engaged and engageable patients encourage more interpretations and are also more likely to improve), these do not operate in as clear a manner with supportive interventions. Support is likely applied more uniformly to all patients, with variation only in response to the needs of the alliance, and does not reflect the overall level of engagement of the patient. Studies that have shown an association between support and outcome either studied a broader range of subjects, included greater variation in therapeutic techniques (Kolden, 1991), collected a more general level of support that overlapped significantly with alliance (Bachelor, 1991; 1994; Orlinsky & Howard, 1967; 1986; 2004), or measured interaction techniques more carefully (Gabbard et al., 1994; Gaston et al., 1994; Gaston et al., 1998; Kernberg et al., 1972; Rounsaville et al., 1987).

Although as shown in Chapter 5, therapist negative feeling about the treatment (i.e., countertransference in the broad sense) is more common in psychoanalysis, where raw and painful affects are more frequently being exposed, we hypothesized and found that such negative feelings did not distinguish between good and bad outcome. This is particularly interesting as the overlap between the improvers and psychoanalytic groups was high, meaning that the finding changed only on the basis of two patients. There has

been support for the common-sense view that a positive therapist feeling would be predictive of good outcome (Rounsaville et al., 1987), but this is typically seen in short but not long treatments (Baer et al., 1980; Saunders, 1999). Any such effect in this study is likely counterbalanced by the complexity of affective interchanges (Hölzer et al., 1997, see below) and the importance of painful affects (on the part of both patient and therapist) in successful analytic and therapeutic work.

6.5.2 Patient variables

Links between patient process variables and outcome are less well represented in the literature and have historically more often been the subject of psychoanalytic research as part of the search for early predictors of analytic success. The findings of this study broadly support the analytic idea that patients who supply material rich in dynamic content in their first year are more likely to go on to symptomatic improvement. Using the patient analogue of Jones operationalisation of dynamic technique, it was found that patients who were described by their analyst as providing dynamic material (including feelings, unconscious content, defences, historical material, transference, in-session behaviour, clear themes, dreams, sex, self-image, and guilt) in the first year, were more likely to be symptomatic improvers. This supports Jones's belief that dynamic content and interventions are predictive of positive outcome across therapeutic modalities (1993; Ablon & Jones, 1998) as well as the more general finding that patients who are "involved" and "participating" in their treatment do better (Baer et al., 1980; Gomes-Schwartz, 1978; O'Malley et al., 1983; Windholz & Silberschatz, 1988).

This study supports the specific prediction from the empirical and theoretical literature that the presence of relationship and transference themes in the patient's material should be associated with positive outcome (Bachrach, 1993; Erle, 1979; Graff & Luborsky, 1977; Norman et al., 1976; Oremland et al., 1975; Orlinsky et al., 1994; Schlessinger & Robbins, 1974). It still remains to establish whether the active discussion

of relationship and transference themes has a causal relationship with outcome or is simply found more frequently in the first year of those patients who go onto improve. Either way, however, the finding suggests an important and clinically useful predictor for who should remain in psychodynamic psychotherapy and psychoanalysis, and who is likely to benefit from it.

The ratio of unconscious content to resistance emerges from the factor analysis in Chapter 4 and from the introduction to this chapter as one of the most natural representations of what psychoanalytic theorists and clinicians see as central to therapeutic action. As predicted, patients who rate higher on therapist-reported unconscious themes and lower on resistance in the first year are more likely to improve symptomatically by termination. In successful treatments, clinicians analyse and reduce the resistance in order to bring important unconscious themes into the content of the session and into the patient's awareness (Fonagy, 1999b; Jones, 1997; Kernberg, 1999). This finding is reminiscent of Samstag's (1998) result that therapist ratings of "smoothness" (i.e., lack of resistance) and patient ratings of "depth" (i.e., important and previously unconscious themes) are associated with positive outcome. The current study demonstrates that therapist ratings of "depth" and "smoothness" are alone valuable in predicting outcome.

Guided by strong non-psychoanalytic findings of the association between a patient's positive view of therapy and a good outcome, this study predicted (and failed to find) such an association. A more careful reading of the literature reveals that this association is particular to short, non-dynamic therapies (Orlinsky et al., 1994; Rounsaville et al., 1987; Saunders, 1999) but does not hold true for the beginnings of longer dynamic therapies (Graff & Luborsky, 1977; Jones & Pulos, 1993). Therapist ratings of patient affect may also be confounded by the therapist's complex impressions of the patient and therefore lack the simple power of patient report. The affective

exchange between therapist and patients is a complicated and important area that requires careful study of reciprocal interactions if it is to yield reliable prediction of patient outcome. Krause's group has used measures of affect to test the theory that successful psychotherapy is characterized by the therapist failing to conform to the patient's maladaptive unconscious role relationship patterns (Anstadt, Merten, Ullrich, & Krause, 1997; Dreher et al., 2001; Krause et al., 1998; Merten, Anstadt, Ullrich, Krause, & Buchheim, 1996; Steimer-Krause, Krause, & Wagner, 1990). In a series of studies applying the EMFACS to videotaped psychodynamic psychotherapy sessions, they found that high prevalence of interactive emotional patterns in the first session was a predictor of negative outcome and that compensatory, rather than reciprocal, affective facial behaviours were indicative of positive outcome (IPA, 2001b). Complementary combinations of predominant affective expressions (called lead-affects), such as an affectively positive patient and negative therapist or negative therapist and positive patient were predictive of the best outcome (Dreher et al., 2001). Over the course of a therapy, changes in affective patterns (e.g., becoming more diversified and increased blending of primary affects) may indicate structural change and can be used to assess the success of the treatment (Dreher et al., 2001).

Bänninger-Huber (1992; 1997) uses the EMFACS to identify "traps" whereby the patient seeks to enlist the therapist in the enactment of a maladaptive interaction pattern (such as the therapist adding to the patient's guilt or legitimizing behaviour to relieve guilt). Within these traps, which last from 15 seconds to 2 minutes, she focuses particularly on "prototypical affective microsequences" (PAMs) which she believes are nonverbal mechanisms for relationship regulation (such as mutual smiling or laughing of patient and therapist) and are central to the success or failure of the traps. Research that identifies PAMs and measures the extent to which the therapist recognizes them and responds in a therapeutic fashion would have clear predictions for therapeutic outcome

and important implications for understanding mechanism of change. Using their text-based measure, Hölzer and colleagues (1997) observed that therapists verbalize more emotions than their patients, that successful therapists use more emotion words than unsuccessful ones, and that by the end of treatment, successful therapists verbalize more emotions in the category of anger.

6.5.3 Therapist-patient interaction variables

The two predicted findings related to patient-therapist interaction yielded interesting and surprising results. The most robust finding in the entire psychotherapy literature, that patient-therapist alliance predicts improvement (Horvath & Symonds, 1991; Horwitz, 1974; 1996a; Kolden, 1996b; 1996; Marmar, Gaston et al., 1989; 1989; Marziali et al., 1999; Orlinsky et al., 1994; Rudolf, 1991), is not supported by this study. This result may be explained in a number of ways. First, similar to the situation with scales of “supportive” and “exploratory” interventions, the therapeutic alliance scale is based on only a single YAWRS item. Direct questions to the therapist about an abstract process probably lack the reliability and validity of other scales used here, and of most alliance measures in the literature that draw from a large number of concrete items. Second, there are several findings in the literature suggesting that therapist report of alliance, particularly in the early stages of a long-term therapy, is less successful than other measures of alliance in predicting improvement (Fenton et al., 2001; Gunderson et al., 1997). Third, particularly in long-term intensive therapies, it is likely that alliance interacts in a complex way with personality pathology, interpersonal problems, and invasive therapist interventions (e.g., transference interpretations) (Ackerman & Hilsenroth, 2003; Bond et al., 1998; Kolden, 1991; Marziali et al., 1999; Sexton, 1993). Alliance measures have at times been applied successfully to long-term psychoanalytic treatments in predicting outcome, but require greater focus on patient-therapist match, multiple sources of data, and concrete questions (Rudolf, 1991; Sashin et al., 1975).

In contrast, patient and analyst discussion of treatment parameters and contract is significantly related to outcome, matching a prediction from the non-analytic literature (Orlinsky et al., 1994). Of note, this scale was derived from two subsection summary scores and numerous concrete items. Although it is not as theoretically interesting as alliance, this finding demonstrates the value of a concrete scale that measures an important way in which therapist and patient engage. It likely reflects an aspect of the therapeutic working alliance more than it does a direct beneficial effect of the discussion of treatment parameters. More work is clearly required on this and other discrete interventions to understand them in relation to the alliance and outcome.

6.5.4 Change over time

The exploratory use of HLM analysis to predict outcome using linear trends in process scales over the first year of treatment yielded unexpected results. None of the three predicted associations were found, while two unpredicted linear trends were detected. This suggests that the technique itself is valid, as there was sufficient power, even in this very small sample, to produce two significant findings, though not in the expected variables. There are several possible ways to account for the lack of replication of Graff's report of increase in transference themes and decrease in resistance over time (Graff & Luborsky, 1977). First, the power of detection for linear trends was likely low given the small sample size and the stringent statistical requirements of the HLM analysis. Second, the time period itself was limited to only one year, whereas some of the treatments lasted many times that. Third, though we have speculated the lack of variation among subjects may have prevented findings from supportive and exploratory interventions, the variation in how therapists rated their patients over time may have been too large to detect the expected temporal pattern. Finally, the high rates of personality pathology in this sample may have resulted in high early scores of

transference material that did not leave room for a significant increase, and high scores of resistance that would not decrease until well into the treatment.

The failure of regression to increase over time, as suggested by the theoretical and clinical literature, suggests that it is also a poor distinguishing factor between improvers and non-improvers. Several possible explanations may be offered. First, the productive increase in regression as part of the therapeutic process may be obscured by unproductive increases in regression in other patients that lead to unresolved transference neurosis and lack of change. Second, the regression scale itself, drawn from a large number of YAWRS items describing regressive behaviours, may be too concrete and measure aspects of pathology which personality-disordered patients bring to their treatments from the start, not just those aspects of behaviour related to engagement in therapy (as would be easier to measure in neurotic patients). Future work is needed to tease out the individual manifestations of regression by patient and see how these change over time in relation to symptomatic improvement.

There has been some good evidence that temporal trends of process variables are useful in understanding the mechanism of change and in predicting outcome. Bucci has shown that movement between phases of the referential process, as measured by her RA and CRA measures, is associated with change in content measures (CCRTs and FRAMES) as well as positive therapeutic outcome (1997b; Bucci, 1998). Ratings of RA have also been useful in empirically guiding the selection of points in a session where central themes are found (Bucci, 1997a). Spence and colleagues (Mayes & Spence, 1994; Spence, 1998) found in a successful case that the analyst was more likely to intervene (i.e., make any verbal utterance) when the co-occurrence rate of transference pronouns (CORtrans) was high. Furthermore this correlation increased over the course of a six year analysis, suggesting that the analyst is increasingly using the patient's language to determine when interventions are appropriate. Spence also hypothesized that the extent

of separation between “you” and “me” (SEPtrans) would be an indicator of the strength of patient-analyst alliance and should be inversely related to number of interventions. Though this was not born out by the data, the inverse association did appear to increase over time. Finally, Jones and Windholz found that the PQS item “therapist identifies a recurrent theme in the patient’s experience or conduct” is significantly related to CORtrans over the course of a treatment (1990), validating the use of this measure. Reproduction of such findings would require considerable modifications in the methodology used in this study. Methods for distinguishing phases of treatment would have to be validated, possibly requiring a study of individual patterns and depth of interpretations, not simply the presence of material as was recorded with the YAWRS. In addition, the changes described may not happen in the first year of treatment.

Serendipitously, two interesting linear trends were detected by the HLM analysis. Supportive interventions increased more over time in the improvers than in the non-improvers, possibly suggesting that such support was necessary in order to maintain an environment in which transference interpretations and analytic process could take place (Gabbard et al., 1994; Horwitz et al., 1996; Kernberg et al., 1972). Further study of the interaction of support, transference interpretations, and alliance would be helpful in better understanding this trend. In addition, dynamic material was found to decrease significantly over time in the sample as a whole. This has several possible interpretations: (1) general resistance may increase over the first year, though not necessarily as a bad prognostic sign, (2) therapists may change how they complete the YAWRS over time, possibly spending less time on minor themes of dynamic material in an effort to complete the questionnaire more quickly, thus causing an artifactual reduction in rating of dynamic content, (3) as the therapy develops, therapists may pursue more specialized themes in the treatment that fit less easily into the generic categories supplied by the YAWRS, and (4) as resistance increases (though this was not measured itself), therapists

may focus more on supportive interventions and manifest material in order to preserve the alliance.

6.5.5 Study weaknesses

The weaknesses of the YAWRS as a measure and the design of this study are reviewed in Chapters 4 and 5. Most importantly, for the hypotheses proposed, the results are limited by the small sample size and the use of a process measure that measures only the therapist's report of patient and therapist content and behaviour. It is notable that despite these limitations, a number of significant findings were detected, suggesting that therapists are highly valuable reporters and the effect sizes of the associations are large enough to be detected even in a small sample. The findings will always be subject to the possible interpretation that they reflect more a therapist's biased and subjective view of a treatment than the truth of what transpires during the sessions. However, in this regard it must be remembered that no therapist report measure has ever before demonstrated an ability to accurately predict who improves in psychodynamic psychotherapy and psychoanalysis. If the therapist's bias is leading to the significant findings it is because they are unconsciously picking up on factors that will influence improvement status or because their biased judgements of the patient are part of the mutative elements of change. We suspect that the latter is partially true, but that therapists are also reliably reporting important elements of real process.

Another fundamental difficulty in the findings presented here and in Chapter 5 is the near complete confound between treatment modality (i.e., psychoanalysis vs. psychodynamic psychotherapy) and membership in the group of symptomatic improvers. Despite this confound, it was possible to ask specific questions of the samples that are relevant to the different comparisons. However, on a whole it is still impossible to know how the modality and improver status were related, and which of the process variables were more tied to one than the other. Future studies must have larger sample sizes and

carefully study unsuccessful psychoanalyses and successful psychodynamic psychotherapies, not just the inverse.

6.5.6 Future research

Future studies of process-outcome links must address a number of the design weaknesses and confounding variables in this study. The YAWRS requires extensive reliability and validity work before it can become a standard, useful measure of process. Measures of process completed by patients and observers must be incorporated into the study design to validate the YAWRS and provide information, particularly with regard to patient affective experience, alliance, and regression, that are least accurately described by therapists. The measures must be applied to a larger sample and with greater consistency over entire courses of psychotherapy and psychoanalysis to collect an adequate data set.

A number of specific subtleties of process material have been exposed in the results, as reported above, that would be valuable foci for individual exploratory studies, and eventually lead to revised process measures. First, as is already being explored in the literature on affect (Anstadt et al., 1997; 1992; Bänninger-Huber & Widmer, 1997; Dreher et al., 2001; Krause et al., 1998; Merten et al., 1996; Steimer-Krause et al., 1990), reliable measures of affective interchange would be useful in quantifying patient and therapist states of mind and relating them to individual mutative therapist interventions, such as transference interpretations. A similar strategy is being applied to a detailed study of language and may help to understand mutative processes in both a micro- and macro-analytic way (1997b; Bucci, 1998; Canfield et al., 1991; Mayes & Spence, 1994; Mergenthaler & Kächele, 1996; Russell & Trull, 1986; Spence, 1998). These methods could be incorporated into detailed analyses of interpretive sequences, making it possible to study the complicated relationship between interpretations, supportive interventions, alliance, and change (Gabbard et al., 1994; Gaston et al., 1994; Gaston et al., 1998; Kernberg et al., 1972; Milbrath et al., 1999; Rounsaville et al., 1987). All of this work

would require multiple sources of information (i.e., patient, therapist, and observer ratings, verbatim transcripts for analysis of speech, and possibly videotapes and/or psychophysiological measures to study affect) and data over long time periods.

Although the categorical measure of symptomatic change used in this study (as described in Chapter 3) has been an improvement on limited or more short-term assessments, it still does not live up to the goal of measuring long-term structural change, an essential part of theorizing about the value of psychoanalysis and psychodynamic psychotherapy. Future studies need to collect data years after therapy has terminated and apply measures that move beyond symptoms and look at structure (as described in Chapter 2). Relation of these measures to process variables will be an important move forward in this work.

Several of the process scales analysed above could be easily parsed further in future work to study the intricacies of their relation to outcome. Supportive interventions take many shapes and should be better understood in their effect on alliance and outcome. It would be particularly interesting to study whether certain kinds of supportive interventions relate most to transference interpretations, or whether the relationship is more by their timing. Transference and relationship material and interpretations all showed interesting findings in the above analyses, but there is insufficient data to understand how relationship and transference discussions relate to one another, and whether they have differential effects in causing or reflecting outcome. Alliance and measures of therapeutic participation (by both patient and therapist) have been studied carefully elsewhere and these methodological advances should be applied to the YAWRS, as well. Likewise there are more careful measures of therapist-patient fit and therapist-patient discussion of important issues, such as the treatment contract. The failure of therapeutic alliance to provide the expected result and the strong finding with discussion of contract both call out for more investigation.

6.6 Conclusion

Application of YAWRS data from the first year of treatment to predict symptomatic outcome, as presented in this chapter, is an exciting step forward in psychoanalytic outcome research. The statistical analyses reported above directly address fundamental theoretical and clinical questions about psychoanalysis and psychotherapy and test previous empirical findings with a novel measure. Data from the YAWRS confirm hypotheses that dynamic therapeutic technique (as specified by Jones, or more specifically by the theoretical constructs of interpretation, transference interpretation, and relationship interpretation) is associated with better outcome. We then added important findings to a sparse literature on the link between patient material and outcome by finding that dynamic material, transference themes, relationship themes, and the ratio of unconscious content to resistance are all predictors of positive outcome. A few other predictions, including the expectation that alliance and linear trends in transference interpretations, resistance, and regression would predict outcome were not born out, most likely due to aspects of the YAWRS and the sample. Unfortunately, the possible confound between treatment intensity and outcome could not be avoided in this small sample. Despite several methodological problems, the data yielded interesting findings that set the stage for future research.

CHAPTER 7. ADULT ATTACHMENT INTERVIEW

7.1 Introduction

In Chapter 1 we reviewed some of the outcome studies of psychoanalysis and psychodynamic psychotherapy that have set out, and partially succeeded, to demonstrate the efficacy of these treatments for different disorders. In Chapter 2 we reviewed the accomplishments and difficulties of psychotherapy research methodology, with a focus on two major challenges: measuring process and capturing structural change, and providing an introduction to attachment theory. Chapter 3 described the Anna Freud Centre Young Adult Research Project, in which 25 subjects were treated with either psychoanalysis or psychodynamic psychotherapy, and change was measured according to standard procedures. In Chapters 4, 5, and 6 we presented data from this study, using a new analyst-report measure of process. In this chapter we will review measures of structural change, describe the relevance of attachment to measurement of psychopathology and therapeutic improvement, and then present data using the Main & Goldwyn Adult Attachment Interview (AAI) to demonstrate that this is a useful measure for process-outcome research.

7.1.1 Measures of structural change

Given the commonly repeated lament that studies of psychoanalysis and psychodynamic psychotherapy are limited by the lack of an effective measure of structural change, there are several surprising observations that emerge from a review of the relevant literature. Rather than there being too few measures of structural change, it is impressive to see how many such measures have been devised in the past quarter century. From the early years of poorly operationalised and unreliable metapsychological checklists (Bellak, Hurvich, & Gediman, 1973; Semrad, Grinspoon, & Fienberg, 1973; Wallerstein, 1986), through a more practical emphasis on psychological capacities (Høglend, 1993a; 1993b; 1993; 1995; 1993; 1992; Horowitz et al., 1986; Kaltreider,

DeWitt, Weiss, & Horowitz, 1981; Sandell, 1987a; Wallerstein, 1988; Weinryb, Rössel, & Åsberg, 1991a; 1991b; Wilczek, Weinryb, Barber, Gustavsson, & Åsberg, 2000), to modern individualized measures of change (Grande et al., 2004; IPA, 2001c; OPD Task Force, 2001; Rudolf, Grande, & Oberbracht, 1997; Sifneos, Apfel, Bassuk, Fishman, & Gill, 1980), object relations measures (Blatt & Auerbach, 2003; Fonagy, 2001d), and case formulation techniques (Barber & Crits-Christoph, 1993; Dahl, 1991; 1998; 1994; Henry, Strupp, Schacht, & Gaston, 1994; 1994; Luborsky, Crits-Christoph, & Mellon, 1986; 1994; Perry, 1989b), there are at least 20 major methodologies, each of which has supported a small literature on reliability, validity, and application to psychotherapy research. The problem appears to have been, however, that outside of briefly citing some of the historical efforts, the designers of new measures were more interested in founding a new methodology than on building carefully from past mistakes or justifying why they had rejected previous methodologies in the first place.

Some researchers have noted that many of the original metapsychological measures, including those of Menninger PRP, Bellak, Semrad, Karush, and May & Dixon, were too abstract, too cumbersome, not sufficiently reliable, and were too highly correlated with symptom measures (Høglend et al., 2000; Mintz, 1981). The PICS (Patterns of Individual Change Scales) failed in that it was not sufficiently sensitive to capture changes during brief psychotherapy with neurotic patients (Horowitz et al., 1986). Individualized measures of change are theoretically attractive, but no one has sufficiently solved the problem of how to standardise the scales such that they can be aggregated in group research. Measures such as the MSI (McGlashan Structured Interview, McGlashan, 1984), SPC (Wallerstein's Scales of Psychological Capacity, Wallerstein, 1988), KAPP (Weinryb's Karolinska Psychodynamic Profile, Weinryb et al., 1991a), CHAP (Sandell's Change After Therapy, Sandell, 1987a), and Dynamic Capacity (Høglend, 1993a; 1995), while promising, have not been used widely enough to

demonstrate their reliability or sensitivity in capturing change. It is notable, however, that the subscales of these measures which achieve the highest reliability scores are those tied to interpersonal behaviours. Not surprisingly, the most promising recent focus has been on measures of interpersonal functioning and object relations. Transference focused measures are attractive, particularly in that they can theoretically be used as measures of both process and outcome. However, to date, the methodologies are still too labor intensive, make too many assumptions about what can be learned from a transcript of a psychotherapy session without a consistent or structured probe, and lack adequate tests of reliability and validity. The efforts to apply these measures to common transcripts, so as to study their similarities and differences is commendable. However, if the efforts of these research groups had been combined to perfect a single such measure, instead of distributing their work among seven similar measures, perhaps we would be further along in our goals.

Given these lessons, a useful measure of structural change for the current study is one that satisfies several criteria: (1) it should have a cohesive theoretical base, preferably one that takes a mainstream psychoanalytic view and incorporates an object relational perspective, (2) it should be sufficiently experience-near to reliably capture operationalisable attributes of a person's behaviour, (3) it should be based on an observer rating scheme, taking advantage of the clinical intuition of such a rater to note unconscious patterns, yet not rely on subtle phenomena that only highly trained raters would observe and may disagree about, (4) it should have an already sizable literature demonstrating that it is reliable and valid in normal and clinical samples. We propose that the Adult Attachment Interview and coding system is the only currently available measure that satisfies these criteria. The theoretical basis for this measure was described in Chapter 2. In this chapter, we will review relevant data regarding its use in measuring psychopathology and therapeutic change.

7.1.2 *Attachment and psychopathology*

The question of how attachment experiences and patterns relate to child and adult psychopathology has permeated the field of attachment from the beginning. As far back as Bowlby's 1960 paper on grief and mourning in infancy, he raised the question of how and when experiences of loss can lead to pathological development and whether they are related to specific types of personality disturbance and psychiatric symptomatology in adulthood (1960). In its earliest formulations, attachment theory postulated that early experiences of caregiving have consequences for relationships and mental health throughout the life cycle. However, it was much less clear to what degree insecure attachment patterns were necessary for psychopathology to occur, how inevitable psychopathology was in a person with an insecure attachment pattern, and whether particular attachment patterns would necessarily lead to specific psychiatric disorders or symptoms. Since then a variety of researchers have formulated more specific theories on how attachment and psychopathology are related, and begun to test these hypotheses in clinical samples.

Main (1996) identified five attachment-related risk factors which influence the development of mental disorders: (a) failure to form an attachment between 6 months and 3 years (maternal deprivation), (b) the "organized" forms of insecure attachment status, (c) major separations from and permanent loss of attachment figures, (d) disorganized attachment in response to early maltreatment, and (e) disorganized attachment as a second-generation effect of the parent's own trauma. She suggested that as our ability to measure states of mind with respect to attachment and their development across the life-cycle improves, we will be able to find more specific links with psychopathology. She also stressed the role of genetic and other factors in the understanding of attachment patterns and psychopathology.

Bowlby and other theorists explained the relationship between attachment patterns and psychopathology by describing the nature of the defensive manoeuvres that an infant undertakes in order to deal with a non-optimal attachment environment, and how those defences affect internal working models and a model of the self (Dozier, Stovall, & Albus, 1999). For example, when infants have experiences which lead them to expect caregivers to be rejecting or undependable, they develop a model of the self as unloved or rejected and a model of the other as unloving or rejecting. If a caregiver is consistently rejecting, the infant is likely to defensively turn attention away from his/her distress and attempt to “minimise” the expression of attachment needs. Alternatively, if a caregiver is inconsistently caring, the infant is more likely to defensively turn their attention toward his/her distress and attempt to “maximise” the expression of attachment needs.

Defensive minimizing is believed to appear in infancy as an insecure-avoidant Strange Situation classification and, barring important mutative life events, in adulthood as a dismissing AAI classification. Dozier suggests that in both childhood and adulthood the minimizing strategy causes the individual to turn away from the self, leaving negative representations unresolved, which then manifest themselves in externalising behaviour. On the other hand, defensive maximizing appears in infancy as an insecure-avoidant Strange Situation classification and in adulthood as an entangled AAI classification. The maximizing strategy predisposes children and adults to internalising symptoms and disorders because attention is riveted on caregiver availability and negative representations remain painfully alive (Dozier et al., 1999). Dozier warns that despite the clarity of this model, it is difficult to make direct predictions from Strange Situation and AAI classification to psychopathology because typically diagnosed disorders are very heterogeneous and comorbid, making the distinction between externalising and internalising manifestations difficult.

Carlson and Sroufe (1995) focus on the importance of the organization function of the attachment system in integrating affective, motivational, and behavioural components of experience. Insecure attachment is believed to predispose to psychopathology because of the disruption it causes to this system. This theory leads to predictions of psychopathology resulting from a disorganized infant attachment classification or a “cannot classify” AAI classification. Crittenden (1995) also emphasizes the organizing function of attachment, specifically its ability to integrate cognition and affect. Avoidant infants become defended against affect and thus depend too heavily on cognition, sometimes developing “false affect.” This is believed to lead to depression with affective distancing, and compulsive behaviour. Resistant infants use a coercive strategy that is based primarily on affect and may develop “false cognition.” This leads to affective displays, impulsive behaviour, agitation, anger, and suspicion. In a disorganized infant, neither affect nor cognition works and a totally disabling depression may be the result (Crittenden, 1995).

Fonagy (1997) takes a different approach to understanding the relationship of attachment and psychopathology. He suggests that sub-optimal infant relationships have the power both to establish maladaptive mental representations and to cause inhibition of mental functions. Most significantly, he argues, insecure infant attachment can lead to an inhibition of mentalizing function, preventing the developing child from adequately seeing the minds of others and, as a result, forming an inadequate representation of the self. This may lead to various forms of psychopathology, and in severe cases borderline personality disorder.

In recent years, dozens of empirical studies have attempted to spell out specific associations between attachment patterns and psychopathology. The two most significant collection of this work have been in a special section of the *Journal of Consulting and Clinical Psychology* (February and April, 1996) and in a recent chapter by Dozier and

colleagues (1999). In both cases, evidence is reviewed linking infant and adult attachment measures with affective disorders, anxiety disorders, dissociative disorders, schizophrenia, borderline personality disorder, and antisocial personality disorder. We will briefly review studies that linked the AAI to several of these disorders.

Affective disorders

Bowlby (1980) described three types of infant attachment experiences that are associated with later depression: loss of an attachment figure, inability to form a secure relationship with caregivers, and a relationship with attachment figures in which the infant is made to feel incompetent or unlovable. Empirical evidence from using the AAI to retrospectively assess childhood experience has begun to confirm this prediction. Fonagy and colleagues (1996) found that subjects diagnosed with depressive disorders scored higher on the AAI probable experience scale for “rejecting” and lower for probable “loving.”

Studies linking depression with specific AAI classifications have been somewhat inconsistent in their findings. Cole-Detke and Kobak (1996) studied a sample of 61 college women and found that subjects with depressive symptoms were more likely to use “hyperactivating” AAI strategies, corresponding with an E classification. They theorized that women who hyperactivate their attachment systems may become overly preoccupied with their own shortcomings, focus excessively on relationships, and fail to develop autonomy and competence, thus leading to depression. The study was complicated by the comorbidity of depression and eating disorders, which may be associated with opposite attachment patterns; however, the finding held even when eating disorder pathology was controlled. Rosenstein and Horowitz (1996) found a similar result in 60 psychiatrically hospitalized adolescents. In this sample affective disorders and dysthymic personality traits (without comorbid conduct disorder) were positively associated with an E classification on the AAI. Finally, Fonagy and colleagues

(1996) found that depressed psychiatric inpatients had significantly higher preoccupied-involved anger and lower idealization (state of mind subscales of the AAI), both of which suggest a more E and less Ds classification for those subjects.

On the other hand, Patrick and colleagues (1994) report that 6 of 12 dysthymic subjects received a Ds classification on the AAI. Aside from the small sample size, this finding is confounded by the fact that dysthymic subjects were selected specifically to be free of any borderline personality symptomatology. Dozier (1999) suggests that differences in attachment findings associated with depression can be attributed to the fact that depression can be either internalising or externalising, and the selection criteria of individual studies may bias a sample to being more one than the other. Fonagy and colleagues (1996) divided depression into major depression, dysthymia, and bipolar disorder and found that major depression was most likely to be associated with an F classification and high anger subscale scores (suggestive of an E classification) and bipolar disorder with a Ds classification. Tyrell and Dozier (1997) also found a high proportion of subjects with major depression to be secure on the AAI. They suggest that this may be due to the episodic nature or greater genetic component of major depression. There is also inconsistency in the link between depression and unresolved status. While Rosenstein (1996) and Patrick (1994) find proportions of U classification in their depressive samples similar to that found in a low risk population (18% and 16%, respectively), Fonagy (1996) and Tyrell (1997) found higher proportions of U in depressive patients (72% and 50%, respectively). This is likely due to the high comorbidity in the latter two studies between depression and personality disorders.

Limited data are available on the link between suicidality and attachment status. Adam and colleagues (1996) report an association between E and U classifications and self-reported suicidal behaviour or severe suicidal ideation in a sample of 133 male and female adolescents. They explain this by theorizing that dismissing subjects set aside

suicidal feelings along with their attachment difficulties. According to Dozier's (1999) theory, suicidality is more of an externalising behaviour and therefore should be associated with Ds qualities. However, as of yet, there is no data to back up this hypothesis.

Anxiety disorders

Bowlby (1973) suggested that almost all anxiety disorders are best accounted for by anxiety regarding the availability of attachment figures. Possible early environments leading to anxiety disorders include both consistently neglecting parents and over-involved or intermittently neglecting parents, theoretically leading to either form of insecure infant attachment. Dozier (1999) divides anxiety disorders into those with primarily internalising symptoms, in which fear predominates (generalised anxiety disorder [GAD], panic disorder), and those with primarily externalising symptoms, in which avoidance predominates (agoraphobia, panic disorder, phobias, and obsessive compulsive disorder). In isolation, the first group of disorders might be expected to be associated with an E classification on the AAI, and the second with a Ds classification, but comorbidity of the two makes distinction difficult.

In a longitudinal attachment study, Warren and colleagues (1997) found that infants with resistant attachments are more likely than infants with secure or avoidant attachments to be diagnosed with anxiety disorders as adolescents. Assuming that the association between infant attachment and adolescent AAI holds, this suggests that an E classification would be associated with anxiety disorders. Rosenstein and Horowitz (1996) support this conclusion, finding that E adolescents had higher anxiety scores on the Millon Personality Inventory (including obsessive compulsive traits). Cassidy (1995) described higher scores on anger and vulnerability (using the Inventory of Adult Attachment) in subjects diagnosed with generalised anxiety disorder. Pianta and colleagues (1996) found that anxiety, as measured by the Minnesota Multiphasic

Personality Inventory (MMPI), was lowest in Ds subjects and highest in E subjects.

Finally, Fonagy (1996) found a higher proportion of unresolved transcripts in subjects with anxiety disorders. This might be interpreted as suggestion that anxiety is either the cause or the consequence of disorganization in response to a traumatic experience.

Eating disorders

Cole-Detke and Kobak (1996) found that when depressive symptomatology was controlled for, eating disorders were significantly associated with a deactivating, or Ds-like, attachment strategy. They explain that a woman with an eating disorder attempts to direct attention away from her own distress by controlling the external world through her body. They also point out that when this strategy is combined with a hyperactivating or E pattern, typical of depressive symptomatology, a CC or U classification may result. Their scoring was done, however, with the Kobak Q-sort and did not use these classifications. Fonagy and colleagues (1996) found a higher proportion of U's among patients with eating disorders, most of whom were also depressed.

Borderline personality disorder

A large proportion (70%-85%) of patients with borderline personality disorder (BPD) are believed to have experienced or witnessed physical or sexual abuse in childhood, raising the strong possibility that attachment processes are involved in the etiology of this condition (Dozier et al., 1999). Fonagy (1997) hypothesizes that such abuse causes an inhibition of the capacity for reflective function in the infant and child, leading to pathological development of self-identity and relationship patterns. This theory is well-supported by Fonagy's (1996) finding that patients with BPD (n=36) scored significantly lower on a reflective function scale, and had lower loving and higher neglecting probable experience scales on the AAI. In addition, 75% of BPD patients were classified as E, and 89% were classified as U. Most strikingly, though, half of the BPD patients who were found to be E were assigned to the rare E3 subgroup, designated

“fearfully preoccupied by traumatic events.” Patrick and colleagues (1994) found that 10 out of 12 female borderline patients were classified as E, all of them E3. Nine of these 12 patients (75%) were found to be U. Rosenstein and Horowitz (1996) also found a significant association between E classification and BPD in psychiatrically hospitalized adolescents. The Cornell Westchester N of 17 and Swiss Grant studies promise to shed more light on this question as they will report AAI classification on a large number of carefully characterized borderline patients. Of the two subjects from the N of 17 study described in the literature to date, one has been said to be preoccupied and the other dismissing (Diamond et al., 1999).

Antisocial personality disorder and conduct disorder

Bowlby (1973) proposed that children may feel intense anger when separated from attachment figures, or when threatened either explicitly or implicitly with abandonment. As this anger may be dangerous in the maintenance of the needed relationship with the parents, it is often repressed and displaced onto other targets. Such displacement can lead to severe behavioural disorders or violence associated with conduct disorder, and later in life, antisocial personality disorder. Fonagy describes a similar pathway from another point of view. He sees a secure attachment as being necessary for development of a mentalizing capacity that inhibits the natural tendency for antisocial behaviour (Fonagy, 1998). Dozier (1999) classifies antisocial and conduct disordered behaviour as externalising disorders that should theoretically be related to a minimizing or Ds-like attachment strategy.

All three theories are supported by the evidence that conduct disorder and antisocial personality disorder are associated with insecure-Ds attachment classifications. Allen and colleagues (1996) found that criminal behaviour in adulthood is predicted by CC and Ds classifications, and higher scores on derogation and lack of resolution with respect to trauma measured 10 years earlier during adolescence. Rosenstein and

Horowitz (1996) found increased incidence of conduct and antisocial personality disorders in Ds adolescents. When unaccompanied by an affective disorder, conduct disorder was not associated with U status; conduct disorder plus an affective disorder was associated with both Ds and U classifications. Lyons-Ruth (1996) showed an association between disorganized attachment in infancy and aggressive behaviour in childhood. Fonagy (1996) found associations between E and U classifications and antisocial personality disorder in psychiatric inpatients. In a sample of prisoners, he found more violent crimes associated with lower reflective function and increased U classification (Fonagy, 1998). An increased incidence of U status in violent men has also been observed in domestic violence research (Holtzworth-Munroe, Stuart, & Hutchinson, 1997).

A significant amount of the ambiguity in relating attachment classification to severe character pathology, such as that of antisocial and borderline personalities, may be related to overtaxing a rating system that was not designed with these populations in mind. Turton and colleagues (2001) list more than a dozen ways in which the current AAI manual does not adequately specify how to rate situations that are far more common in patients with “extremes of experience” (such as in forensic samples) than in the general population. These include lack of a clear set of attachment figures, numerous losses, extreme violations of discourse, generalised derogation of attachment, self-derogation, culturally appropriate idealisation, institutional care, and psychiatric illnesses that include cognitive damage or thought disorders. The authors suggest general guidelines for deciding when such transcripts are rateable and advocate for further research in these populations (Turton et al., 2001).

Several studies have demonstrated that like conduct disordered and antisocial behaviour, substance abuse and dependency are associated with a Ds attachment classification on the AAI. Allen (1996) found that lack of idealization, derogation, and

involved anger predict hard drug use in adults. Rosenstein and Horowitz (1996) showed that a Ds classification is associated with substance abuse in adolescents. Mickelson and colleagues (1997) also related an avoidant attachment pattern (as measured by Hazan and Shaver's attachment self-report measure) to alcohol and drug abuse.

Despite the fact that many studies have shown links between insecure subclassification on the AAI and specific types of psychopathology, some researchers and theorists have concluded that these associations are weak at best and that the most important finding is the general relationship between insecurity and psychopathology. Allen and colleagues (1996) found a predominance of insecure AAI classification in a sample of adults who were psychiatrically hospitalized as adolescents, but no association between the type of insecurity and the adolescent diagnosis. Reviewing the findings on attachment and psychopathology in children and adolescents, del Carmen (1996) concludes that more precise diagnostic instruments are needed to make clear links between attachment processes and psychiatric diagnosis. Scott Brown and colleagues (2003) reported greater interpersonal difficulties and psychiatric symptoms in adolescents with ambivalent attachment patterns, as compared to secure or avoidant subjects. They reasoned that a hyperactivating strategy in an adolescent is likely to lead to a range of psychological difficulties.

Several other researchers point out that attachment and psychiatric diagnosis are difficult to relate because current DSM-III-R or DSM-IV diagnoses pay little attention to attachment related factors (Dozier et al., 1999). Mickelson and colleagues (1997) concluded that other than alcohol and drug abuse, there is little association between psychopathology and a subtype of insecure attachment. In a major meta-analysis of current normative data in attachment, van IJzendoorn and Bakermans-Kranenburg (1996) concluded that clinical samples are overwhelmingly insecure, but exhibit a wide range of classifications that do not appear linked to individual diagnoses. Although

Fonagy and colleagues (1996) reported several associations between AAI subtype and DSM-III-R diagnosis, the vast majority of significant findings reported in their study served only to establish that severely disturbed psychiatric patients are overwhelmingly insecure, as measured by classification and AAI subscales, when compared with a normal low risk sample.

Several researchers have made the prediction, based on theoretical models and descriptions of E and Ds transcripts, that subjects with an E classification are more likely to self-report psychiatric symptoms than Ds subjects, even when controlling for an objective measure of symptomatology (Dozier et al., 1999). Pianta (1996) administered the AAI and MMPI to 110 first-time mothers and found that E subjects were highest on several psychiatric indices of distress and relationship problems, and Ds subjects were lowest on the same measures (with F subjects somewhere in the middle).

At this stage in research on attachment and psychopathology it is possible to make a number of general observations about the relationship of AAI classification and psychopathology and point out the obstacles that have kept investigators from establishing more definite links. Without exception studies of clinical samples have found that psychiatric inpatients and outpatients have higher rates of insecurity and lack of resolution with respect to trauma and loss (U status) than non-clinical subjects. It is important to point out, though, that other than Warren's (1997) observation of the link between resistant attachment and anxiety disorders in adolescence, these studies are cross-sectional and thus do not offer any information as to the nature of this association. At least four possibilities are apparent: (a) insecure attachment is a risk factor for the development of psychopathology (Fonagy, 1998; Slade, 1999), (b) psychopathology is a risk factor for development of insecure attachment, (c) psychopathology makes a subject appear insecure on the AAI, (d) some independent agent is a predisposing factor for

both psychopathology and insecure attachment. All four include a wide range of possible mechanisms that will require a great deal of future research if they are to be worked out.

Several confounding and confusing factors must be addressed in order to better understand the association between attachment and psychopathology. First, theorists and experimentalists disagree whether attachment strategies are best understood according to discrete categories or continuous scales. Advocates of discrete groups must define strict and reproducible measurements of severity of psychopathology and demonstrate that these scales are significantly higher in an insecure attachment category. Advocates of continuous scales should be able to show significant correlations between attachment scales and severity of psychopathology. In either system, a greater uniformity of standards for classifications and scales must be established across the literature. With respect to the AAI, this implies use of only certified coders. Second, an attempt must be made to isolate the existence of psychopathology from other typical concomitants of psychiatric disorder, such as negative life experiences, low socioeconomic status, and psychiatric care. For example, it is possible that an insecure attachment is more related to the experience of being under psychiatric care than to psychopathology itself, because many studies of psychopathology have focused on samples who are receiving some sort of treatment.

Although not unequivocal, evidence does support Dozier's theory that a deactivating, Ds-like, attachment strategy is associated with externalising symptoms, while a hyperactivating, E-like, attachment pattern is associated with internalising symptoms. Dozier is correct to point out the heterogeneity of psychiatric disorders; the distinction between internalising and externalising symptoms in these disorders must be more carefully operationalised before more conclusive studies can be done.

Unfortunately, few studies of attachment and psychopathology have available the large and diverse samples that are needed to disentangle the effects of comorbidity and

heterogeneity of disorders. Finally, more studies will need to assess psychopathology from a variety of viewpoints (e.g., self-report, family-report, clinician-report, objective assessment of behaviour) in order to explicate the relationship between attachment and reporting of psychopathology, as well as to actual psychopathology in different domains.

6.1.3 Attachment and psychotherapy

Although research on attachment and psychotherapy has grown significantly in the past few years, it is somewhat disappointing that this work was not begun earlier, given that Bowlby's theory of attachment was intended to have most influence in this area. This fact is likely due to a combination of factors which caused Bowlby and his work to be excluded from mainstream psychoanalysis for many years and Bowlby's own relative weakness in theorizing about the psychotherapeutic process (Slade, 1999). More recently, the application of attachment theory to clinical work may have been hampered by the mistaken assumption that attachment theory is always consistent with the work of the British objects relation school and therefore has little to add to that already large literature (Slade, 1999). Nonetheless, this gap has begun to be addressed in the last decade, when many prominent theorists and researchers have turned their attention to how attachment theory helps with a formulation of the therapeutic relationship and process (Diamond et al., 1999; Eagle, 1997; 1999b; Fonagy, 1999c; Holmes, 1997; 1998; Lichtenberg, 2003; Muscetta, Dazzi, De Coro, Ortu, & Speranza, 1999; Slade, 1999) and what attachment measures can tell us about the process and outcome of therapy in children and adults (Dozier, Cue, & Barnett, 1994; 1996; Fonagy et al., 1995; Hardy et al., 1999; Hesse, 1999; L. M. Horowitz et al., 1993; Juffer, van IJzendoorn, & Bakermans-Kranenburg, 1997).

When applying attachment theory to the therapeutic process, most writers are careful to point out that despite the apparent links between insecurity and psychopathology, the two are not synonymous and attachment organization is better

thought of as a risk factor or typical relationship pattern, than as the pathology that psychotherapy sets out to repair (Slade, 1999). Nevertheless, some theorists draw links between theories about attachment in the therapeutic process, and existing theories about specific psychopathologies and how they may be treated. Slade (1999) associates a dismissing state of mind with respect to attachment with obsessional, schizoid, and narcissistic character pathologies and a preoccupied/entangled state of mind with hysterical or borderline personality disorders. In his extensive work on the role of reflective function in psychopathology and psychotherapy, Fonagy (1999c) applies similar theories to borderline personality disorder and to patients with E and U attachment classifications.

The danger in this approach is that just as therapists are wary of narrow diagnostic categories when referring to the complex personality factors and diverse defence mechanisms in their patients, they are reluctant to base their thinking on a set of four or five discrete attachment categories. This is particularly true during psychotherapy, when defences may have become quite fluid and a patient is likely to fluctuate among multiple modes of attachment and defence. Slade (1999) suggests emphasizing the regulatory strategies associated with a given attachment classification, as opposed to the classification itself. She argues that inherent in each insecure category are aspects of other insecure categories, which may ordinarily exist out of conscious awareness but are accessible in the clinical situation. For application to the clinical setting, it may be better if attachment theory and research focus more on continuous scales (such as the subscales of the AAI), or at least on multiple or repeated measures of attachment classification.

Primarily, most attachment theorists believe that a major goal of psychotherapy is “the reappraisal of inadequate and outdated working models of self in relation to attachment figures” (Bretherton, 1995). Holmes (1998, p. 236) describes the goals as a “search for intimacy and autonomy”; therapy aims “to provide an environment that

fosters attunement, is secure enough to cope with relevant protest, and therefore can allow new meanings and secure-autonomous narratives to arise.” He also divides the work of psychotherapy into story-making (helping the patient tell a coherent story) and story-breaking (allowing the story to be told in a different, more healing light). In these terms, dismissing patients tend to cling to rigid stories, preoccupied patients tended to be overwhelmed by “unstoried” experience, and unresolved patients are unable to find a narrative strong enough to contain traumatic pain (Slade, 1999).

Fonagy (1999b) divides the therapeutic process into three potential and possibly concurrent phases: (1) allowing the patient to externalise alien parts onto the therapist so as to make therapy tolerable, (2) helping the patient disinhibit mental processes (chiefly reflective function) by focusing on mental states in the context of an attachment relationship, and (3) fostering the reorganization or restructuring of the patient’s representational system through interpretation of transference in the here and now. Attachment theory plays an important role in understanding phases 2 and 3 of this model. During phase 2, reflective function, highly correlated with coherence of mind and transcript on the AAI and an inherent part of the secure/autonomous attachment classification, must improve in order for therapy to ultimately be successful. In phase 3, internal working models of self, other, and self in relation to other are altered by changes in procedural memory (Clyman, 1991; Crittenden, 1995).

The principle distinction between different applications of attachment theory to the psychotherapeutic process lies in the extent to which therapy is a “cognitive effort towards understanding of the unsatisfactory working model” or a “fundamental attempt to offer recapitulation through a positive attachment experience” (Mace & Margison, 1997). Bowlby and Holmes tend to favor the first approach, while Fonagy, Main, and Slade tend towards the second.

In all theories of attachment and psychotherapy, attachment is believed to have an especially important role in describing the therapeutic relationship, particularly the transferential and countertransferential aspects of this bond. It stands to reason that the nature of the patient-therapist bond is influenced by the states of mind with respect to attachment of both therapist and patient. The challenges and opportunities that the bond poses for psychotherapy should be intimately related to attachment classification.

Empirical studies designed to explain the relationship between measurements of attachment and psychotherapeutic process and outcome have begun to appear in the last 10 years, and are beginning to report their findings. These studies can be divided into four categories: (1) client attachment pattern as a predictor of therapeutic process, (2) client attachment pattern as a predictor of therapeutic outcome, (3) therapist attachment pattern as a predictor of therapeutic process, (4) change in client attachment pattern as a measure of treatment outcome. The results of these studies will be explored in relation to predictions from attachment theorists.

Several attachment theorists have proposed that the security of a client, in their own state of mind with respect to attachment and in their relationship with the therapist, is an important factor in a successful therapy (Main, 1995; Slade, 1999). Slade (1999) suggests that patients respond to therapists according to their lifelong patterns of defence and affect regulation and therefore are more likely to engage in the work of therapy if they have a secure classification. Main (1995) proposes that fluidity of attention, capacity for reflection on mental states of self and others (metacognitive monitoring), and the ability to establish relatively unambivalent relationships are also features of a secure individual that favor the success of therapy. Dismissing and preoccupied patients are more likely to experience substantial anxiety in conjunction with efforts to alter an enduring pattern of attachment. Dismissing patients will cling to old patterns and block affective experience and memory via “nodal memories” (rigid, inflexible versions of the story,

Holmes, 1998). E patients will present to therapy with a superficial understanding of relationships and confusing, vague, and overwhelming feelings, which need to be contained by the therapist (Holmes, 1998?; Main, 1995). Unresolved patients, particularly if their attachment classification is secure, are more likely to hide central difficulties during therapy and require the therapist to actively pursue exploration of traumatic experiences (Main, 1995). Such exploration may be slow and painstaking and may engender terror and dissociation (Liotti, 1995).

The evidence is strong that dismissing patients pose a distinct set of challenges to therapists. Dozier (1990) found that Ds individuals appear more resistant to treatment than secure individuals on a variety of measures. They deny help, reject treatment, and try to divert the clinician's attention when emotional issues are finally confronted. It is less clear whether this presentation is always accompanied by the underlying attachment issues believed to characterize Ds patients. Hardy and colleagues (Hardy et al., 1999) selected 10 psychotherapy sessions rated as most helpful by clients, and classified them, applying principles from Main's AAI coding manual, according to client attachment styles, presenting attachment issue, and therapist responsiveness. Clients with dismissing attachment styles did not necessarily present with issues about abandonment, abandonment, rejection, and being alone as the authors expected, but did present with issues about loss and proximity seeking. Clients with preoccupied attachment styles did not necessarily present with anger, bitterness, conflict, danger, a need to be close, a need to be cared for, and proximity seeking. These findings are limited, however, by the accuracy of assessing attachment style or presenting attachment issues from a single session. Hardy and colleagues did find that therapists responded to clients according to the rated attachment style: therapists responded more to a preoccupied style with reflection and to a dismissing style with interpretation (Hardy et al., 1999).

Liotti (1995) described interpersonal behaviour in children with a disorganized attachment pattern within the therapeutic setting that matched a theoretical understanding of this pattern: disorientation and distancing in the dialogue, and expression of multiple and incoherent attitudes toward the therapist. Korfmacher and colleagues (1997) found that mothers responded to home visit and group therapy interventions for their conduct-disordered children (?) according to their attachment classifications. Unresolved mothers had the greatest difficulties with the facilitators of the interventions and were most likely to adopt a “crisis orientation” to the therapy. Dismissing mothers were most emotionally shallow and least engaging in supportive therapy, and secure mothers showed the greatest emotional commitment and participation. Dazzi and colleagues (Muschetta et al., 1999) have developed a coding system for therapeutic sessions, identifying Ds-like and E-like violations of Grice’s maxims from session transcripts. They found that coherence violations decrease as sessions progress (Hesse, 1999).

It is straightforward to extend existing theories and research on the association of client attachment classification and the client’s behaviour in therapy (i.e., the transference) to predictions as to how a client’s attachment style affects the therapist’s behaviour (i.e., the countertransference). One might expect therapists of dismissing patient’s to feel “caught in a barren landscape”, hopeless about achieving change or attaining intimacy, angry, unacknowledged, or inept (Slade, 1999). Just as it does in other relationships, the dismissing style excludes the therapist, and leaves him/her with a sense of being intrusive, melodramatic, and even ridiculous. The therapist may respond by unintentionally forgetting things that he/she meant to bring to the patient’s attention or collaborating with the patient in the avoidance of transference issues. Alternatively, a therapist may experience a more sadistic reaction, attempting to force the dismissing

patient to acknowledge disturbing feelings in part out of frustration and feeling rejected by the patient (Slade, 1999).

A therapist's countertransference to a preoccupied-entangled patient may be just as important, but of a very different character. In response to the patient's hyperactivating style the therapist may feel swamped, confused, and disregulated and fear that collaboration is impossible. The therapist may become overzealous in his/her attempts to compose a coherent story for a patient who is not able to make use of such a story. Progress is more likely to follow when the therapist shows that they are able to tolerate the patient's fragmentation and chaos and maintain long-term emotional availability (Slade, 1999).

Although it stands to reason that a therapist's reaction to a patient is associated with attachment classification, it is also possible that the fluidity of defence mechanisms and attachment experiences achieved during therapy makes the countertransference difficult to predict from only an initial assessment of attachment classification. An initially dismissing style may give way, on occasion, to overwhelming affect that produces a countertransference more akin to that expected with an entangled patient. On the other hand, a therapist may feel that a patient's entangled style serves to keep them from forming a lasting intimate relationship with the therapist, and thus feel rejected and excluded – what one more readily associates with a dismissing patient. To date, no research has looked at this question and it remains to be seen whether or not there is a reliable relationship between attachment style and countertransference.

Data on the relationship between client attachment style and therapeutic outcome is scarce and somewhat mixed. In the only large scale study of this question, Fonagy and colleagues (1996) assessed 82 patients in a psychodynamically-oriented psychiatric hospital on the AAI and the Global Assessment of Functioning (GAF) scale of the DSM-III-R at the beginning and end of therapy (mean length of stay = 9 months). They

found that Ds patients were more likely to improve on the GAF than either E or F patients (rates of significant improvement = 93%, 41%, and 33%, respectively). DSM-III-R diagnoses, a battery of psychometric measures assessing depression, anxiety, and personality traits, and the U status of the patients on the AAI were not predictive of treatment outcome. In a preliminary analysis of data used in the study presented in this thesis, Fonagy and Tallandini (1993) found that patients classified as preoccupied-entangled at initial assessment were most likely to prematurely drop out of treatment. Fonagy speculated that directing the attention of a Ds patient to those issues that have previously been avoided is a smaller barrier to overcome than recasting a well-engrained set of attachment beliefs in an E patient. In addition, the perpetual activation of the attachment system in E adults may interfere with the formation of a therapeutic relationship.

On the other hand, Horowitz and colleagues (1993) reported that patients with a dismissing style of describing their parents (as measured by the inability of a naïve listener to reliably recall elements of their description) were less likely to improve in response to brief psychodynamic psychotherapy. The relevance of this finding for understanding the predictive value of attachment in psychotherapy is limited. First, “dismissing” subjects were compared with those less “dismissing”, but not necessarily more preoccupied, as in Fonagy’s study. Second, the recall measure used is likely to be more a measure of coherence (i.e., ability to give a description that someone else can understand and remember) than of a dismissing attachment pattern as measured by the AAI. Finally, the findings are limited to brief psychodynamic therapy, which may be particularly poorly suited for Ds patients because of the limited amount of time in which to form a relationship.

If, as current psychotherapy theory suggests, the therapy relationship is as dependent on the therapist as it is on the client (Beutler et al., 2004), it should be

expected that the therapist's state of mind with respect to attachment is also a critical factor in determining therapeutic process. Bowlby (1988) suggested that the therapist must be sensitive, empathic, and secure in order for the patient to separate childhood projections from real experience in psychotherapy. Recent attachment theorists have postulated the existence of a caregiving behavioural system, distinct but in parallel with the attachment behavioural system, that may account for the ability of therapist to provide the a corrective environment for a patient (George & Solomon, 1999). Main (1995) predicted that therapists classified as F on the AAI are most likely to be able to serve as "secure bases" for their patients and engender successful outcomes. She speculated, however, that therapeutic training, therapeutic process, or the recognized necessity of assisting others in distress may permit some insecure therapists to override their insecure state and provide good caregiving to their patients. She describes three insecure mothers of infants who had recently suffered a life-threatening illness, all of whom were evaluated as securely-attached on the Strange Situation. Main postulated that the outer-directed attention required by a mother of a sick child, perhaps akin to that required of a therapist with a needy patient, overrode the mother's insecure attachment and allowed a positive response to the infant.

An empirical study of this question (Dozier et al., 1994) found that an interaction of client and social worker attachment styles, as measured by the AAI and Kobak Q-sort, is a good predictor of depth of social worker intervention and perception of the client. Insecure social workers were more likely to see their preoccupied clients as highly dependent and intervene more intensely with them than their dismissing clients. This suggests that they were less able than secure social workers to hear and respond to the needs of dismissing patients, manage the demands of the preoccupied patients, and use their own countertransference to respond to subtle manifestations of need and dependency.

A growing area of attachment and psychotherapy research has sought to identify a link between client and therapist attachment classifications and the development of a therapeutic alliance (Mallinckrodt, 2000). Rubino and colleagues (2000) noted that anxiously attached therapists produced less empathic responses to video-taped mock patients, particularly when the patients were fearful or secure. Eames and Roth (2000) found lower scores on alliance between experienced therapists and fearfully attached clients. Interestingly, Tyrell and colleagues (1999) found that clients who had a more deactivating attachment style formed better alliances with therapists who were less deactivating, and vice versa. Dozier and colleagues (2001) reported that clients with dismissing strategies showed less manifest rejection of their case managers than did preoccupied clients, but made their significant others feel worse and had more trouble making use of appointments with case managers. This suggests that though dismissing patients may superficially appear to be better patients they ultimately are harder to help. Finally, Sauer and colleagues (2003) found that anxiously attached therapists initially appeared to form better alliances with their patients (as reported by the patient), though this pattern was reversed as the treatment progressed.

One of the most exciting prospects for the application of attachment theory and research to psychotherapy is the use of attachment measures as indicators of improvement in response to treatment. Psychotherapy efficacy research has continuously faced the problem that available measures, particularly those that assess primarily psychiatric symptoms and diagnoses, do not capture the sometimes subtle intrapsychic and structural changes that psychotherapy purports to achieve. This is particularly a problem for researchers seeking to demonstrate that psychodynamic and/or long-term psychotherapies are preferable to more superficial and/or short-term approaches (see Chapter 2). Zilberg, Wallerstein and colleagues (1991) suggest that psychotherapy researchers need measures of structural change that are identifiable with specific

personality theories and assesses stable psychological capacities “needed to achieve adaptive functioning and life satisfaction.” Measures of adult attachment fit these requirements well: they are clearly identifiable with a theoretical model of development and measure a usually stable capacity that it linked to adaptive functioning.

At the most basic level, it would be expected from attachment theory that successful psychotherapy is associated with a shift from insecure to secure states of mind with respect to attachment, increased coherence of “stories” regarding attachment relationships, and increased resolution with respect to any experienced losses and traumas. Fonagy (1995) reported that of 35 patients beginning intensive inpatient psychiatric treatment, all of whom were initially rated as insecure on the AAI, 40% were rated as secure on discharge (mean duration of treatment = 9.4 mos). Ratings of probable experience scales did not shift significantly between the two assessments, but there was a slight tendency for judges to rate the patient’s experience as less loving and more neglecting and rejecting on discharge. This may indicate that patients were able to communicate the difficult experiences of their childhood more clearly after treatment. As part of the Cornell N of 17 study, Diamond and colleagues (1999) report a change from insecure (one E and one Ds) to secure classification in two borderline patients undergoing psychodynamic psychotherapy. Interestingly, in only one of these cases (the E patient) is the change in classification associated with evidence that the patient improved their object representations. Lichtenberg (2003) questions the relationship between a new attachment classification and structural change in this study, cautioning that after only one year of treatment any such change may be more akin to “transference cure” than to a lasting shift in character. Further data on attachment classification is pending from this study and a larger randomized trial of psychodynamic psychotherapy.

A few other studies report the use of attachment as an outcome measure. Travis and colleagues (2001) found that 29 clients entering time-limited dynamic psychotherapy

were all rated as insecure initially, whereas 7 of them were rated as secure after a mean of 21 sessions. Mean security in the group also improved. Muscetta and colleagues (1999) report gradually increasing attachment security in the case study of a dissociating adolescent patient, which they associate with his symptomatic improvement.

Several studies have used attachment measures as indicators of change in response to other types of intervention. Erickson, Kormacher, and Egeland (1992) used the AAI to measure change in a mother's internal representations following an intervention designed to improve sensitivity to her child's needs. Although separate measures of sensitivity did show change, the mother's attachment classification and the infant's security with respect to her did not. In a case study, Juffer and colleagues (1997) administered the AAI to a mother before and after a four-session intervention designed to improve her sensitivity. The mother was classified as Ds3 both before and after the intervention; however, her sensitivity rating did improve and the infant seemed more secure with her.

Far too few studies of attachment measures as indicators of psychotherapeutic change have been reported to make any definite conclusions. At one extreme, Fox (1995) suggests that the finding that mothers do not seem to change their AAI classification in response to an intervention program may be an indicator that the AAI is not truly measuring the constructs it claims to be. He is particularly critical of Egeland's suggestion that if repeating the study measuring maternal response to intervention, he would have the intervention focus more on her attachment status and feelings towards her own parents. How could he ever claim, Fox asks, that the AAI is reflecting a change in attachment organization crucial to producing a secure infant if the intervention is specifically designed to encourage the discourse measured in the AAI to sound coherent?

A more moderate explanation of the lack of a clear relationship between AAI change and psychotherapy success may relate to how central a role internal working

models play in psychotherapy. Harris (1997) proposes that psychotherapy must also be concerned with the external world that a patient returns to. Even if a patient's insecure internal working models are replaced with secure ones, when the therapist does not pay sufficient attention to the "outer context" the patient may return to a depressogenic situation in which relapse is probable. Harris also stresses the importance of therapy being long-term enough for the patient to find a secure-base outside of therapy before he/she loses the secure-base of the therapist. No doubt, length and intensity of intervention are also important variables in predicting whether a psychotherapy intervention will produce change in attachment status. The complexity of factors requires a great many more studies to be done before the answer to these questions are known.

7.2 Hypotheses

Hypothesis 1: The Adult Attachment Interview (AAI) will be used as an outcome measure in a psychotherapy process-outcome study and will be collected at initial assessment, 18-month intervals during treatment, termination, and at 18-month intervals after termination.

Hypothesis 2: At initial assessment patients will appear predominantly insecure (Dismissing or Preoccupied-entangled) and at termination will be increasingly secure, both by AAI classification and subscales (Diamond et al., 1999; Fonagy et al., 1995).

Hypothesis 3: No significant relationships will be found between demographic variables (gender, age, socioeconomic status, and IQ) and (a) initial AAI classification and subscales or (b) change in AAI classification and subscales over the course of the treatment (Bakermans-Kranenburg & van IJzendoorn, 1993; Sagi et al., 1994).

Hypothesis 4: Significant associations will be found between initial psychiatric assessment and AAI variables:

- (a) Subjects with a higher number of personality disorders will have lower positive SOM scale scores (Dozier et al., 1999; Fonagy et al., 1996; van IJzendoorn & Bakermans-Kranenburg, 1996).
- (b) Subjects with borderline, narcissistic, or self-defeating personality disorders will be more likely to have a Preoccupied-entangled classification (Fonagy et al., 1996; Patrick et al., 1994; Rosenstein & Horowitz, 1996).
- (c) Subjects with paranoid, antisocial, or avoidant personality disorders will be more likely to have a Dismissing classification (Allen et al., 1996; Rosenstein & Horowitz, 1996).
- (d) Severity of initial depressive and anxiety symptoms will not be associated with any AAI classification or subscales.

Hypothesis 5: Subjects in psychoanalysis will be more likely to change their AAI classification and subscales in the direction of higher security than subjects in psychodynamic psychotherapy (Fonagy et al., 1995).

Hypothesis 6: Significant associations will be found between symptomatic improvement status and AAI variables:

- (a) Initial AAI classification will not be associated with symptomatic improvement (Dozier, 1990; Fonagy et al., 1996).
- (b) Symptomatic improvement will be associated with change in AAI classification and subscales in the direction of higher security (Diamond et al., 1999; Fonagy et al., 1995).

Hypothesis 7: Follow-along analysis will show that patients in psychoanalysis become more insecure during treatment, but by termination are less insecure than at initial assessment.

Hypothesis 8: Follow-up analysis will show that changes in AAI classification or subscales are maintained 18 months after termination.

Hypothesis 9: Limited associations will be found between the YAWRS (described in Chapter 4) and AAI classifications and scales, based on the small subset of subjects with both types of data.

- (a) No meaningful associations will be found between YAWRS scales and initial AAI classification or initial AAI subscales
- (b) Higher scores on dynamic technique, dynamic material, and ratio of unconscious themes to resistance will be positive associated with a move towards security on AAI classification and subscales.
- (c) No association will be found between supportive interventions and change in AAI classification and subscales.

7.3 Methods

7.3.1 Administration and coding of the AAI interview

The AAI was administered to all the subjects in the current study, along with other psychological and psychiatric assessment measures, at the time of initial assessment (usually before treatment had begun), at approximate 18 month intervals during treatment, at treatment termination and at 18 month intervals after termination. The interview was administered by two psychiatrists, each trained in administration but not scoring of the AAI. Anna Higgitt administered all interviews to subjects in intensive treatment, while Anthony Bateman administered all interviews to non-intensive subjects. Interviews took between 45 and 90 minutes to complete and were audio recorded using a Sony dictating tape recorder.

In total, 79 AAIs were conducted on the 25 subjects in the study, ranging between one and five interviews per subject. In four interviews tape recording equipment failed and the resulting tape could not be transcribed. On the 75 remaining interviews, transcribers trained according to Main's protocol (Main, 1991), produced verbatim

transcripts of the recording. These transcripts were verified for accuracy by another listener who removed references to whether the interview was an initial or follow-up assessment, and when during the course of the treatment the interview had taken place. Finally, they were distributed to eight different certified AAI raters, all of whom had attended an AAI institute led by Main and Hesse, had adequate agreement at the time of coding on at least 18 of Main and Hesse's reliability cases, and were either already certified or went on within a year to complete the 30 cases necessary for certification. Cases were distributed such that no coder rated the same subject more than once and coders received a random distribution of transcripts from intensive and non-intensive subjects, and from various times before, during, or after treatment. No AAI's were coded until most treatments had already ended, ensuring that coders did not necessarily code earlier assessments first and had no way of knowing at what point within a treatment the interview had been administered.

Coders assigned transcripts four- and five-category classifications (excluding and including U status, respectively), as well as a full set of experience and state of mind scales, as defined in the AAI coding manual. One coder found two transcripts to be uncodable because of protocol violations in AAI administration. On 18 of the 73 coded transcripts, judges indicated that it was not possible to assess the unresolved status of the subject because probing of losses and traumatic events had been insufficient.

7.3.2 Sample and other measures

Sample characteristics have been described in Chapter 3. For the purposes of testing hypotheses, AAI variables were compared with demographics, treatment parameters, initial assessments, and classification of change as described in Chapter 3. For the final set of statistical analyses, AAI variables were compared with YAWRS scales, as described in Chapter 4 and used in Chapters 5 and 6.

7.3.3 *Statistical analysis*

For the purposes of statistical analysis, AAI classification was condensed to a three-way system of Dismissing (Ds), Preoccupied-entangled (E), and Secure (F). Because of the scarcity of CC codes (only two at initial assessment), such transcripts were re-assigned the next best classification. The classification of “Unresolved with respect to trauma or loss” (U) was not used in any analyses because errors in interview administration made these classifications unreliable. The AAI probable experience (PE) subscales were prepared for data analysis by averaging mother and father scores, yielding five subscales: loving, rejecting, involving/reversing, pressure to achieve, neglecting. AAI state of mind (SOM) subscales were analysed both individually (averaging coherence of transcript and mind into a single scale), yielding six subscales (idealizing, involving anger, derogation, metacognitive processes, passivity, coherence) and averaging all seven subscales together in the appropriate direction, yielding a single “Positive SOM” subscale (= mean of 9-idealizing, 9-involving anger, metacognitive processes, 9-passivity, coherence of transcript, and coherence of mind).

For analyzing change in AAI classification, the 16 subjects with initial and termination AAIs were subdivided into four groups: (1) F to E, (2) F to F, (3) insecure to insecure (E to E, Ds to Ds, or Ds to E), and (4) Ds to F. Change in AAI subscales was studied by calculating a pre-post score, subtracting initial from termination scores for each scale and subject. Though this method is not ideal because of its vulnerability to regression to the mean effects and its questionable clinical meaning, it was used because the Jacobson-Truax and other more sophisticated means of evaluating change require psychometric statistics which are not available for the AAI subscales.

Data analysis was conducted using the Chi-squared, ANOVA, and Pearson correlation routines of SPSS 10.0, as appropriate to the comparison being made. Missing value substitution was unnecessary given that statistics were performed in a univariate

fashion, and the only missing data consisted of AAI subscales which appeared to be randomly distributed. Because of the exploratory nature of the data analysis and the specificity of individual hypotheses, Bonferroni corrections were not used and α was set at 0.05 for all analyses.

7.4 Results

7.4.1 Summary

A summary of AAI classifications by assessment point is presented in Tables 7.1a and 7.1b. As described in Chapter 3, due to difficulties in recruiting and scheduling subjects to return for follow-along and follow-up appointments, many assessments were not administered at the proper time. Combined with four equipment failures and two uncodable interviews, this led to a much smaller pool of AAIs than intended. Of the 25 subjects in the study, 24 had successful initial AAIs, 16 had both initial and termination AAIs, 8 had initial, follow-along, and termination, AAIs, and only 7 had initial, termination, and follow-up AAIs. Full AAI classification data on each of the intensive and non-intensive subjects is presented in Appendix 7.1. A summary of AAI subscales by assessment point is presented in Tables 7.2a and 7.2b. Of the 24 subjects for whom initial AAI assessment was available, five were Ds (three intensive, two non-intensive), six were E (one intensive, five non-intensive), and 13 F (nine intensive, four non-intensive). Of the 16 subjects for whom both initial and termination AAI assessments were available, four were Ds (two intensive, two non-intensive), three were E (one intensive, two non-intensive), and nine were F (seven intensive, two non-intensive). Of the Ds subjects, one remained Ds (intensive), one became E (intensive), and two became F (both non-intensive). Of the three E subjects, all remained E. Of the nine F subjects, six remained F (four intensive, two non-intensive) and three became E (all intensive).

Eight subjects, all in psychoanalysis, had data available for follow-along AAI assessments. Two subjects (E and X) were classified F consistently at initial, follow-along, and termination. One subject (U) had a single classification of E at one of two follow-along points, but otherwise remained F consistently. Three subjects (A, C, and D) began as F's, made the transition to E during follow-along, and remained E at termination. One subject (P) was classified E consistently at initial, follow-along, and termination. Finally, one subject (T) was classified as Ds at initial and termination, but as E at follow-along.

Follow-up data were available for seven subjects (three intensive, four non-intensive). Three subjects (H, L, and R, all non-intensive) were secure at termination and remained secure at follow-up. Two subjects (M and P, both intensive) went from E at termination to F at follow-up. One subject (Q, intensive) went from F at termination to E at follow-up and one subject (F, non-intensive) was E at termination and remained E at follow-up.

Tables 7.4a and 7.4b contain mean values for AAI subscales and initial to termination subscale change (positive = increase in subscale) for the five probable experience scales, six state of mind scales, and one state of mind summary scale.

<i>Subject</i> (n=13)	<i>Axis II</i> (n=13)	<i>Initial</i> (n=13)	<i>Follow-along</i> (n=8)	<i>Termination</i> (n=10)	<i>Follow-up</i> (n=3)
A	B	F	E, E, E	E	
B	None	Ds			
C	B	F	E, Ds, E	E	
D	B	F	F, E, E	E	
E	B	F	F, F, F	F	
I	B	F			
M	A/C	Ds		E	F
P	B	E	E, E	E	F
Q	None	F		F	E
T	A/C	Ds	E	Ds	
U	B	F	E, F	F	
V	None	F			
X	B	F	F	F	

Table 7.1a. AAI classifications for intensive subjects.

<i>Subject</i> (n=11)	<i>Axis II</i> (n=10)	<i>Initial</i> (n=11)	<i>Follow-along</i> (n=0)	<i>Termination</i> (n=6)	<i>Follow-up</i> (n=4)
F	B	E		E	E
G	B	E			
H	B	F		F	F
J		E			
K	A/C	F			
L	A/C	Ds		F	F
N	A/C	E		E	
R	A/C	Ds		F	F
S	A/C	F			
W	B	E			
Y	A/C	F		F	

Table 7.1b. AAI classifications for non-intensive subjects.

<i>Subject</i>	<i>RF</i>	<i>Loving</i>	<i>Rejecting</i>	<i>Involving</i>	<i>Press to ach</i>	<i>Neglecting</i>
A	3.5 → 3.5	2.3 → 1	6.5 → 1	3 → 7	1 → 2.5	5 → 1
B	4 →	2 →	7 →	1 →		1 →
C	4 → 2	6 → 1.3	1 → 2.5	3 → 3.5	1 → 1	3 → 1
D	3 →	1.5 → 1.5	4.5 → 3	5 → 6	1 → 2	6.8 → 4
E	2.5 → 4	7 → 2.5	1.5 → 3	3.3 → 4.5	→ 2	3.3 → 2
I	4 →	1 →	7.5 →	2.5 →		7 →
M	2 →	3 → 2.3	1 → 4.5	3 → 4.3	1 → 2	1 → 6.5
P	3 →	3 → 2	4 → 7	2 → 2.8	6 → 2.3	1 →
Q	5 → 5	1 → 2	7 → 5	2 → 3	7 → 4	5 → 4.5
T	2 →	1.3 → 1	7 → 7	2 → 2.5	1 → 1.5	6.5 → 1
U	4 →	7 → 4	2.5 → 2	2 → 3.5	3 → 4	1 → 1
V	3 →	5.3 →	7 →	2 →	1 →	8 →
X	5 → 4	5 → 4	1 → 3	5 → 2.5	2.5 → 1	→ 2

Table 7.2a. AAI RF and probable experience subscales for intensive subjects at intake and termination.

<i>Subject</i>	<i>RF</i>	<i>Loving</i>	<i>Rejecting</i>	<i>Involving</i>	<i>Press to ach</i>	<i>Neglecting</i>
F	1.5 → 3	1 → 4	3.8 → 2	4 → 2	1 → 1	→ 1
G	5 →			9 →	1 →	8 →
H	3 → 6	3 → 5.5	2 → 1	5.3 → 3	1 → 1	6 → 2
J	6 →	2.3 →		8 →	1 →	3 →
K	5 →	4.5 →	3 →	2.5 →	1 →	3 →
L	3 →	4.5 → 4	4.5 → 5	1 → 2	1 → 1	3.5 → 6
N	7 → 4	1 → 1	9 → 9	1 → 1	9 → 9	1 → 7
R	7 → 7	5.5 → 3.5	3.5 → 4.5	3 → 2	→ 1	3.5 → 5
S	4 →	3.5 →	5 →	1 →	1 →	4 →
W	-1 →	1 →	4 →	5 →	1 →	
Y	5 →	4 → 3	3 → 4	1 → 1	2 → 1	2 → 3

Table 7.2b. AAI RF and probable experience subscales for non-intensive subjects at intake and termination.

<i>Subject</i>	<i>Idealizing</i>	<i>Anger</i>	<i>Derogation</i>	<i>Metacog</i>	<i>Passivity</i>	<i>Coherence</i>
A	2 → 2	1 → 3.3	2 → 1	1 → 1	3 → 3	5 → 2
B	4 →	1 →	1 →	3 →	2.5 →	3 →
C	3.5 → 2	1 → 3.5	1 → 1	2 → 1	3.5 → 2	5 → 3
D	1 → 1	4 → 4	2 → 1	2 → 1	3 → 5	4 → 1.8
E	2 → 2.5	1 → 1	1 → 1	3 → 1	1 → 3	8 → 5.8
I	1 →	1 →	1.5 →	3 →	1.5 →	7 →
M	4 → 1.5	1 → 2	3 → 1	3 → 1	3 → 2.5	3 → 4.5
P	2 → 1	2.5 → 4	3 → 1	2 → 1.5	6 → 7	3.5 → 3
Q	1 → 1.5	1 → 1.5	1 → 2	3 → 6	3 → 2	7 → 6.5
T	2.3 → 1.8	1 → 1	2 → 6.5	1 → 1	2.5 → 3.5	4 → 2
U	1.5 → 3	1.5 → 2	1.5 → 1	1 → 4	2.5 → 3.5	5.5 → 4.8
V	4 →	1 →	1 →	1 →	4.5 →	5 →
X	1 → 3	1 → 1.5	1 → 3	3 → 4	3 → 2	0 → 6

Table 7.3a. AAI probable experience subscales for intensive subjects at intake and termination.

<i>Subject</i>	<i>Idealizing</i>	<i>Anger</i>	<i>Derogation</i>	<i>Metacog</i>	<i>Passivity</i>	<i>Coherence</i>
F	1 → 1	1.8 → 2.5	2.5 → 3	1 → 3	6.5 → 7	2.3 → 3.5
G	2 →	4 →	5 →	2 →	5 →	2.5 →
H	3 → 2	2 → 1	1 → 1	1 → 3	3 → 3	5 → 7
J	2 →	5 →	1 →	1 →	3 →	3 →
K	2 →	1 →	1 →	1 →	2.5 →	5 →
L	1 → 1	1 → 2	1 → 1	2 → 1	1.5 → 4	4.8 → 6
N	1 → 1.5	7 → 3	1 → 1	1 → 4	6 → 5	1.3 → 3.5
R	4 → 2	1 → 1	1 → 1	1 → 6	1 → 1	4 → 7
S	2.5 →	1 →	1 →	1 →	3 →	6 →
W	2.5 →	3 →	5 →	5 →		1.3 →
Y	3 → 3.8	1 → 1	3 → 1	5 → 1	1.5 → 3	5.8 → 3

Table 7.3b. AAI probable experience subscales for non-intensive subjects at intake and termination.

<i>AAI Subscale</i>	<i>Initial mean (SE)</i>		<i>Change mean (SE)</i>	
	<i>Intensive (n=13)</i>	<i>Non-intensive (n=11)</i>	<i>Intensive (n=13)</i>	<i>Non-intensive (n=11)</i>
Probable Experience				
Loving	3.48 (0.63)	3.03 (0.50)	-1.55 (0.53)	0.33 (0.60)
Rejecting	4.42 (0.75)	4.19 (0.61)	0.20 (0.75)	-0.04 (0.34)
Involving/reversing	2.75 (0.33)	3.70 (0.86)	0.93 (0.43)	-0.71 (0.38)
Pressure to achieve	2.45 (0.63)	1.90 (0.76)	-0.36 (0.53)	-0.20 (0.13)
Neglecting	4.04 (0.74)	3.78 (0.63)	-1.31 (0.91)	1.40 (1.08)
State of Mind				
Idealizing	2.25 (0.34)	2.18 (0.29)	-0.10 (0.37)	-0.29 (0.31)
Involving anger	1.38 (0.25)	2.52 (0.61)	0.88 (0.26)	-0.54 (0.55)
Derogation	1.62 (0.21)	2.05 (0.49)	0.10 (0.55)	-0.25 (0.27)
Metacognitive processes	2.15 (0.25)	1.91 (0.48)	0.05 (0.50)	1.17 (0.96)
Passivity	3.00 (0.34)	3.30 (0.58)	0.30 (0.35)	0.58 (0.37)
Coherence	4.62 (0.58)	3.70 (0.52)	-0.58 (0.73)	1.17 (0.61)
Positive State of Mind	5.74 (0.23)	5.04 (0.30)	-0.51 (0.18)	0.57 (0.38)

Table 7.4. AAI subscale initial means and change means.

7.4.2 Demographics

No significant association was found between any of the demographic variables and initial AAI classification. Of 48 individual statistics testing the association of demographics and initial AAI subscales, three significant results and one trend were found. Subjects with higher IQ's were found to have lower rejecting ($r(df=11)=-0.71$, $p<0.01$) and higher involving/reversing ($r(df=11)=0.64$, $p<0.02$) scores. Subjects with higher SES were found to have higher involving anger ($r(df=22)=0.52$, $p<0.01$), and women had slightly higher scores on neglecting ($r(df=19)=0.42$, $p<0.06$). No significant association was found between demographic variables and change in AAI classification. One of 48 tests of the association between demographics and change in AAI subscales

was significant: subjects with higher SES were found to have a greater increase in positive SOM over the course of treatment ($r(df=14)=0.54, p<0.03$).

7.4.3 *Baseline assessments of psychopathology*

Associations between baseline measures of psychopathology and initial AAI classification were complex. No straightforward association between continuous psychopathology variables (extent of depression, extent of anxiety, number of DSM Axis I disorders, and number of DSM Axis II disorders) and AAI classification was found. However, when subjects were subdivided on the basis of individual Axis I and II diagnoses, three patterns emerged: (1) All seven patients who were diagnosed on the SADS-L at initial assessment with antisocial, violent, or conduct-disordered behaviour (either current or retrospectively) were rated as Secure at initial assessment ($\chi^2(df=2)=7.06, p < 0.03$). (2) Of the five patients classified as Dismissing at initial assessment, four received a diagnosis of avoidant or paranoid personality disorder at initial assessment (though three patients with these diagnoses were rated Secure and one rated Preoccupied-entangled). (3) Of the four patient classified as Preoccupied-entangled at initial assessment, three received a diagnosis of borderline or self-defeating personality disorder at some point during their treatment (though an additional nine patients with these diagnoses were classified as Secure). Tables 7.1a and 7.1b list the Axis II diagnosis for each of the 24 patients (Cluster B = borderline or self-defeating personality disorder, Cluster A/C = avoidant or paranoid personality disorder).

Eight of 48 correlations between initial AAI subscales and baseline psychopathology were significant or near significant. Number of Axis II diagnoses was positively correlated with rejecting ($r(df=19)=0.54, p<0.02$), pressure to achieve ($r(df=17)=0.40, p<0.10$), and involving anger ($r(df=20)=0.54, p<0.01$) and negatively correlated with loving ($r(df=19)=-0.46, p<0.04$) and positive state of mind ($r(df=20)=-0.41, p<0.06$). Number of Axis I diagnoses was negatively correlated with coherence

($r(df=22)=-0.40, p<0.06$). BDI depression was positively correlated with neglecting ($r(df=19)=0.38, p<0.10$) and STAI trait anxiety was positively correlated with derogation ($r(df=22)=0.36, p<0.09$).

No associations were found between baseline psychopathology and change in AAI classification. Only 1 of 48 correlations between baseline psychopathology and AAI subscale change was significant: there was a positive association between number of Axis I diagnoses and decrease in the passivity subscale ($r(df=14)=0.52, p<0.04$).

7.4.4 Treatment intensity

Associations between treatment intensity and initial AAI classification and subscales were explored to detect possible confounding factors when comparing AAI change in the two groups. Though the association between initial AAI classification and treatment intensity was not significant, it was notable that of the five subjects who received a Preoccupied-entangled initial classification, four were assigned to the non-intensive treatment group. In addition, trends were found towards involving anger being higher ($F(1,22)=3.4, p<0.09$) and positive state of mind being lower ($F(1,22)=3.5, p<0.08$) in the non-intensive group.

A near significant association was detected between change in AAI classification and treatment intensity ($\chi^2(df=2)=5.14, p<0.08$). All three subjects who changed from a Secure to a Preoccupied-entangled classification underwent intensive treatment. Both subjects who changed from a Dismissing to a Secure classification were in non-intensive treatment. Subjects who either remained Secure or Insecure during the course of treatment were evenly divided between intensive and non-intensive groups. Two significant and two near significant associations between treatment intensity and subscale change were found: Loving ($F(1,14)=3.5, p<0.09$) and positive state of mind ($F(1,14)=5.3, p<0.04$) were both found to decrease in subjects in psychoanalysis and increase in subjects in psychodynamic psychotherapy. At the same time,

involving/reversing ($F(1,14)=4.6, p<0.05$) and involving anger ($F(1,14)=4.3, p<0.06$) were found to increase in psychoanalysis and decrease in psychodynamic psychotherapy.

7.4.5 Overall improvement status

The relationships between overall improvement status and AAI classification did not reach significance but showed interesting and important patterns. Of the five initially Dismissing patients, only two went on to become improvers (whereas two out of four Preoccupied-entangled and 7 out of 10 Secure patients became improvers). This is severely confounded by the fact that four out of the five Dismissing patients (including two of the three non-improvers) were in non-intensive treatment (which has been generally associated with less symptomatic improvement, as discussed in Chapter 3). No association was observed between initial AAI subscales and overall improvement status.

Two interesting findings emerged from the relationship between change in AAI and overall improvement. The two subjects who went from Dismissing to Secure (L and R, both of whom were in non-intensive treatment) were non-improvers, whereas only 3 of the remaining 14 subjects were non-improvers ($\chi^2(df=1)=3.81, p<0.06$). Of the 12 AAI subscales tested, only one showed a change difference across improvement groups: passivity increased in non-improvers but remained stable in improvers ($F(1,14)=5.81, p<0.05$).

7.4.6 Follow-along descriptive analysis

Due to the small number of subjects with follow-along AAI data, statistical analysis was not practical on this sample. Figure 7.1 shows the mean positive state of mind scale across five time points (Initial, Follow along 1, 2, and 3, and Termination) for seven intensive improvers and for the single intensive non-improver for which follow-along data were available (subject D). In the intensive improvers, positive state of mind decreased after initial assessment and then rebounded by termination, though was still almost two standard errors below the initial mean. The single intensive non-improver

showed an initial increase in positive state of mind, followed by a rapid decrease and no improvement by the end of treatment.

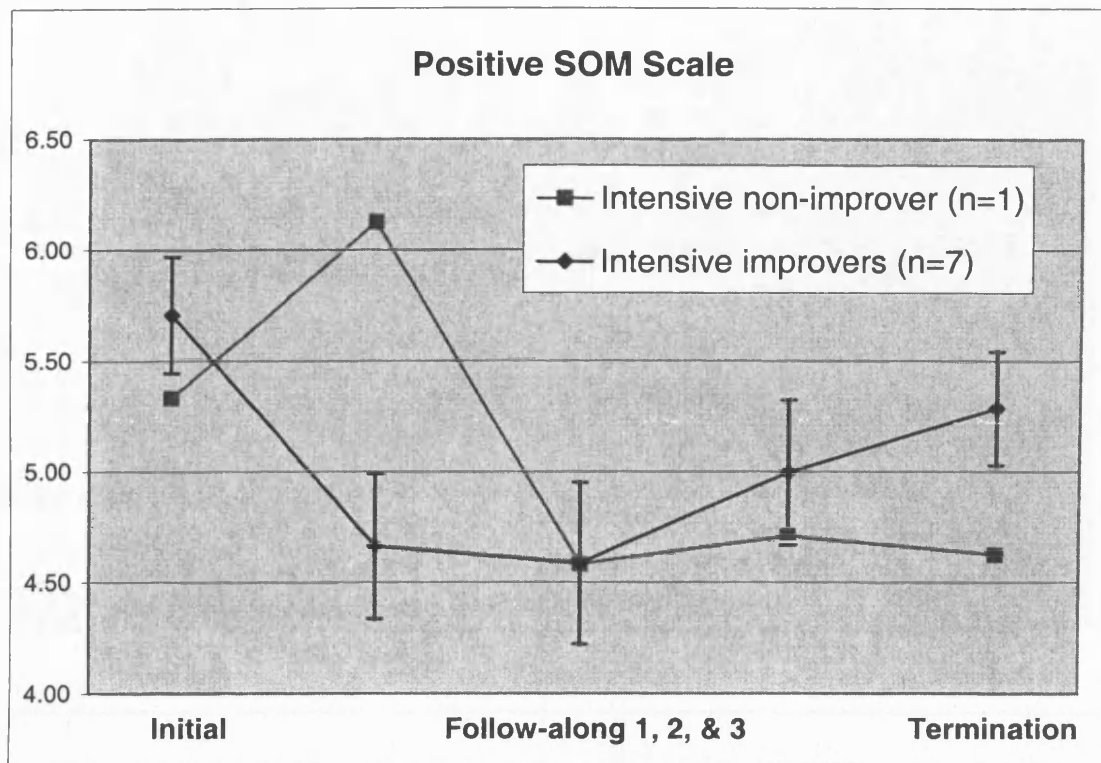


Figure 7.1. Positive SOM scale in follow-along analysis.

7.4.7 Follow-up descriptive analysis

As with the follow-along data, statistical analysis was not practical on the small number of subjects for whom follow-up AAI data were available. Figure 7.2 shows the mean positive state of mind scales for intensive improvers (n=3), non-intensive non-improvers (n=3), and a single non-intensive improver. No significant differences between the pattern of change in these groups were observed.

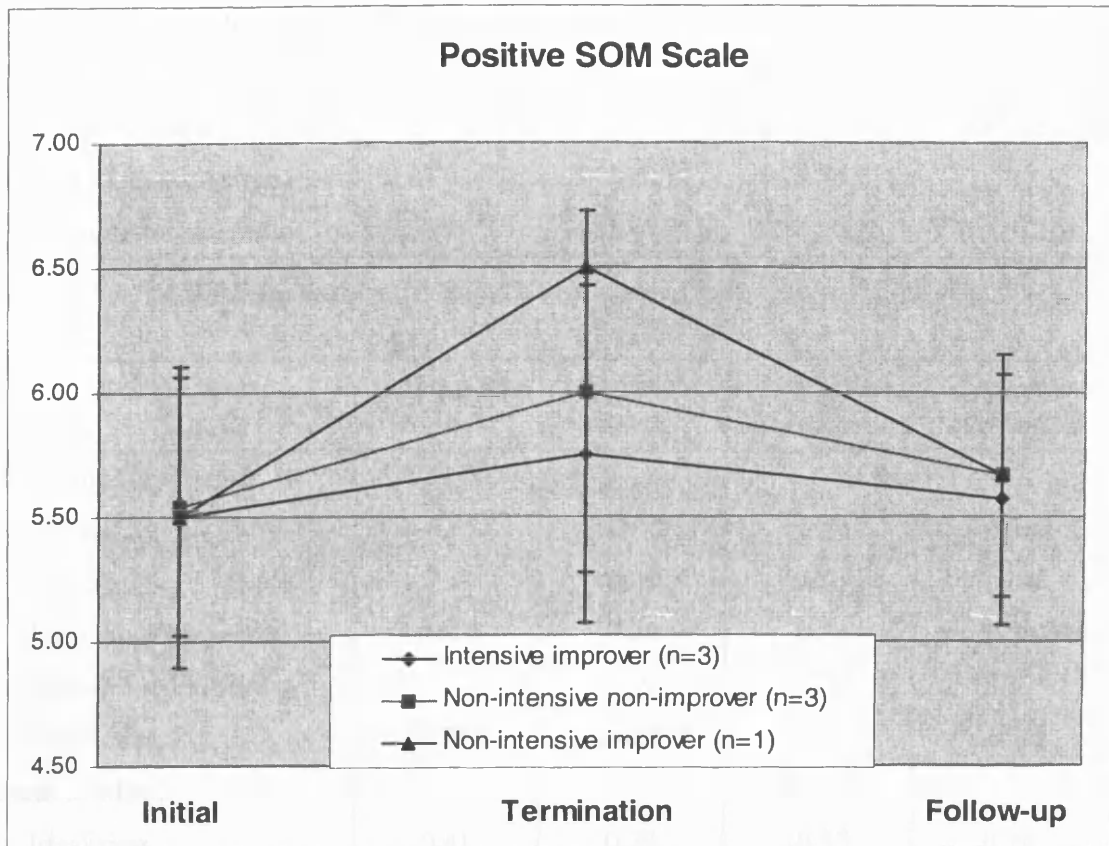


Figure 7.2. Positive SOM scale in follow-up analysis.

7.4.8 AAI and YAWRS

Due to the small number of subjects in this analysis, YAWRS factors were limited to the four most meaningful YAWRS factors described in Chapters 4 through 6: Jones dynamic technique, Jones dynamic material, ratio of unconscious themes to resistance, and supportive interventions. As predicted, the ANOVAs looking for an association between initial AAI classification (Ds, E, or F) and the four YAWRS factors yielded non-significant results (see Table 7.5). Contrary to prediction, 7 of the 44 correlation coefficients calculated between YAWRS factors and 11 initial AAI subscales were significant: the ratio of Ucs to resistance was positively associated with Pressure to Achieve and Anger, and dynamic material and technique were both negatively associated with Neglecting and the composite positive State of Mind scale (dynamic technique was negatively associated with Coherence, as well) (See Table 7.6).

Mean (SD)	Ds (n=4)	E (n=2)	F (n=3)	F (df)
Dynamic technique	0.69 (0.24)	0.76 (0.29)	0.56 (0.06)	0.58 (2,6)
Dynamic material	0.67 (0.23)	0.76 (0.20)	0.60 (0.11)	0.43 (2,6)
Ratio of Ucs to resistance	0.63 (0.03)	0.90 (0.21)	0.68 (0.27)	1.6 (2,6)
Supportive interventions	0.59 (0.59)	0.63 (0.71)	0.31 (0.26)	0.32 (2,6)

Table 7.5. YAWRS factors by Initial AAI classification.

<i>AAI Subscale</i>	<i>Dynamic technique</i>	<i>Dynamic material</i>	<i>Ratio of Ucs to resistance</i>	<i>Supportive interventions</i>
Probable Experience				
Loving	-0.51	-0.40	-0.11	-0.002
Rejecting	0.37	0.26	0.40	0.04
Involving/reversing	-0.10	0.04	-0.11	0.24
Pressure to achieve	0.62	0.59	0.77*	0.59
Neglecting	-0.75*	-0.85**	-0.55	-0.25
State of Mind				
Idealizing	0.41	0.29	-0.53	0.28
Involving anger	0.53	0.52	0.74*	0.46
Derogation	0.15	0.16	-0.12	-0.08
Metacognitive processes	0.16	-0.02	-0.43	-0.34
Passivity	0.36	0.45	0.57	0.14
Coherence	-0.67*	-0.63	-0.41	-0.44
Positive State of Mind	-0.67*	-0.68*	-0.56	-0.52
Reflective function	0.21	0.10	0.14	0.42

* = $p < 0.05$, ** = $p < 0.01$

Table 7.6. Correlations between YAWRS factors and Initial AAI subscales.

No association was found between AAI change category and the 4 YAWRS factors tested (see Table 7.7). However, 4 of the 44 correlations between change in AAI subscales and YAWRS factors were significant (see Table 7.8). Two of these (positive association between decrease in Passivity and dynamic material and technique) matched predictions. In addition, associations were found between supportive interventions and a move toward security on Passivity and overall State of Mind scales.

Mean (SD)	F→F (n=2)	non-F→non-F (n=3)	Ds→F (n=2)	F (df)
Dynamic technique	0.60 (2.4)	0.82 (0.22)	0.51 (0.20)	1.8 (2,4)
Dynamic material	0.64 (0.12)	0.81 (0.16)	0.50 (0.17)	2.4 (2,4)
Ratio of Ucs to resistance	0.78 (0.29)	0.82 (0.20)	0.63 (0.02)	0.51 (2,4)
Supportive interventions	0.29 (0.36)	0.79 (0.57)	0.54 (0.76)	0.44 (2,4)

Table 7.7. YAWRS factors by AAI change classification.

<i>AAI Subscale (positive = decrease)</i>	<i>Dynamic technique</i>	<i>Dynamic material</i>	<i>Ratio of Ucs to resistance</i>	<i>Supportive interventions</i>
Probable Experience				
Loving	0.01	0.06	0.05	0.31
Rejecting	-0.45	-0.29	0.44	-0.46
Involving/reversing	-0.08	-0.15	-0.18	-0.11
Pressure to achieve	-0.26	-0.50	-0.41	-0.57
Neglecting	-0.76	-0.62	-0.24	-0.58
State of Mind				
Idealizing	0.34	0.21	-0.51	0.55
Involving anger	0.61	0.51	0.63	0.46
Derogation	0.33	0.24	-0.39	0.06
Metacognitive processes	-0.08	-0.22	-0.53	-0.47
Passivity	0.91**	0.90**	0.42	0.87*
Coherence	-0.29	-0.25	-0.15	-0.61
Positive State of Mind	-0.54	-0.52	-0.35	-0.83*
Reflective function				

* = $p < 0.05$, ** = $p < 0.01$

Table 7.8. Correlations between YAWRS factors and change in AAI subscales.

7.5 Discussion

7.5.1 Summary

The AAI was collected, as planned, at initial assessment, termination, and for small subsamples, at 18 month follow-along and follow-up periods. Sufficient data were collected for use as both a predictor of treatment outcome and, for 16 subjects, a

measure of structural change. Unfortunately, follow-along and follow-up data were too limited for any formal statistical analyses. This points to the great difficulty and expense required in consistently collecting, transcribing, and coding AAIs. Psychotherapy process-outcome research has historically been plagued by the manpower required by its measures and been weakened by the small sample sizes and incomplete data that are the result of this load. Sadly, this study maintains that tradition and the AAI does not help this problem. However, the large existing database of AAI data, complete with good reliability and validity data, along with the theoretical strength of the measure hopefully counteracts its methodological and practical burden.

Based on the existing literature linking psychopathology and insecurity on the AAI (Dozier et al., 1999; Fonagy et al., 1996; van IJzendoorn & Bakermans-Kranenburg, 1996), the distribution of initial AAI classifications was unusual because of the large proportion (13 out of 24 = 54%) of subjects classified as secure. This contrasts sharply with the 17% average rate of security in 18 samples (drawn from five studies) reported by Dozier and colleagues (1999). Of note, the two highest rates of security in this review were in a small sample of outpatients with unipolar depression [five out of six secure; Tyrell, 1997] and in an inpatient sample of patients with antisocial personality disorder (8 out of 22 secure; Fonagy et al., 1996). Though the Young Adult Subjects were quite ill, as evidenced by their high rates of depression and anxiety, described in Chapter 3, they were all outpatients and were predominantly troubled by depression and cluster B personality disorders. Both of these may be consistent with an initial presentation of superficial interpersonal functioning and attachment security. The approximately equal numbers of Dismissing and Preoccupied-entangled subjects were consistent with other literature on mixed clinical samples (Dozier et al., 1999). The fact that Entangled subjects were more heavily represented in the non-intensive than the intensive treatment group was worrisome for the interpretability of later findings. This presumably was the result of

chance as subjects were assigned alternatively to treatment groups by an administrator who had no knowledge of their diagnosis or route of presentation. However it still has serious consequences confounding the interpretation of differences between treatment groups and initial AAI classification.

The pattern of change in AAI classification was also surprising given the literature-based expectation that subjects in psychotherapy will move in the direction of security. In fact, of the seven initially Insecure subjects for whom termination data were available, only one was classified as Secure at termination, and three of the initially Secure subjects (all receiving intensive treatment) were classified as Preoccupied-entangled at termination. Such changes have never been reported in the literature and require a modification of the currently simplistic view of attachment security as a proxy for therapeutic improvement. However, Lichtenberg (2003), in his critique of papers in three special issues of *Psychoanalytic Inquiry* predicting and reporting a move towards security in small samples of psychotherapy patients, foresaw the intricate nature of this relationship and set the theoretical stage for results such as ours:

For therapists, the dialectic between coherence and incoherence is complex. A move from coherence toward a degree of incoherence may signal the opening up to exploration of an area of affectively loaded struggle. Alternative a move from incoherence to coherence can signify positive change as noted by Muscetta (1999). However, premature restoration of coherence can indicate shutting off the opportunity to explore, while too great a loss of coherence may be the beginning of a period of disorganization (2003, p. 202).

We propose that in dynamic psychotherapy, particularly when it is frequent and intense as in psychoanalysis, part of the core process is for patients to exhibit “confused, angry, or passive preoccupation with attachment figures” exactly as is described by Main in her depiction of a subject with Preoccupied-entangled attachment classification (George et al., 1996). In patients with severe character pathology who at first appear to be Secure, the transition to a Preoccupied-entangled state may be necessary (but not

sufficient) for their successful treatment. At or beyond termination of the treatment, we might then expect the patient to transition back to a Secure classification as evidence of their structural change. Some suggestion of this pattern is seen in the limited follow-along and follow-up data in this study. Five of the eight subjects for whom follow-along data were available showed some evidence of a transition to a Preoccupied-entangled status during the course of treatment (A, C, D, T, and U; three remained E at termination, one returned to F, and one returned to Ds) and four of these were symptomatic improvers. Of the three subjects who were E at termination and for whom follow-up data were available, two were classified as F at follow-up and both were improvers. The one subject (F) who was E at termination and remained so at follow-up was a non-improver (though subject Q, who went from F at termination to E at follow-up was an improver).

AAI subscale initial and change scores were difficult to interpret in absence of a comparison group, as normative data have not been published. Several differences in change scores were observed between intensive and non-intensive groups and will be discussed below under treatment intensity.

7.5.2 Demographics

As hypothesized, no consistent associations were found between demographic variables (gender, age, IQ, and socioeconomic status) and initial AAI classification, initial subscales, change in AAI classification, or change in subscales. Of the 96 correlations tested with AAI subscales and change in subscales, only four were significant or near significant, well within the number predicted by chance. This confirms the well-documented finding that AAI is independent of IQ, age, and other social factors (Bakermans-Kranenburg & van IJzendoorn, 1993; Sagi et al., 1994) and is important for the validity of AAI as a measure of structural change. However, the usefulness of this

result is limited by the small sample size in this analysis, the relatively small range of demographics in the study sample, and the limited number of demographics tested.

7.5.3 Baseline assessment of psychopathology

Despite the complexity of the associations between initial AAI and psychopathology, the results roughly corresponded with expectations based on the available literature. In keeping with the inconsistent and weak findings, found in the literature, on the relationship of depression, anxiety, and other Axis I disorders to AAI classification no such relationships were found. However, similar to Fonagy's (1996) association between depression and negative probable experience scales, a positive correlation was found between the BDI and neglecting. The majority of significant findings centred on the well documented relationships between personality disorders and AAI classifications and subscales. As predicted, number of DSM Axis II disorders was negatively correlated with positive state of mind (Dozier et al., 1999; Fonagy et al., 1996; van IJzendoorn & Bakermans-Kranenburg, 1996). Not surprisingly, given that many personality disorders are associated with early histories of adversity, number of Axis II disorders were also associated with several negative probable experience scales.

The relationship between cluster B personality disorders and the Preoccupied-entangled classification did appear, but was obscured by the fact that such a large percentage of subjects, including those with personality disorders, were classified initially as Secure. This may be due to the fact that these subjects were higher functioning and had had less previous treatment than those studied by Fonagy (1996), Patrick, (1994), and Rosenstein (1996). The AAI was not designed as a clinical instrument and in such subjects confusingly classifies them as Secure until the pathology of their underlying structures is exposed by psychotherapy. Similarly, the expected association between paranoid, antisocial, and avoidant personality disorders and the Dismissing classification was found but was obscured by the unexpectedly high rating of Secure. This was most

striking in the finding that all subjects with antisocial, violent, or conduct-disordered histories were initially rated as Secure. Despite the attachment experiences that were presumably part of the development of these behavioural disorders, these subjects managed to function at a reasonable level by superficially representing their early experiences in a coherent enough way to merit a secure classification. This contrasts with the theories of Fonagy (1996), Dozier (1999), Rosenstein (1996), and Allen (1996) and requires a more subtle understanding of the AAI as applied to outpatient clinical populations. Besides the small sample size, the greatest weakness of the findings associating personality disorder and AAI classification is the absence of data on Unresolved classification. This classification has been shown to be extensively related to personality disorders and may help explain the unexpectedly high incidence of security in this sample (Dozier et al., 1999; Fonagy et al., 1996; Patrick et al., 1994). The lack of probing for resolution with regards to trauma and loss in the AAI interviews may itself also be related to the underclassification of insecurity.

7.5.4 Treatment intensity

Associations between treatment intensity and change in AAI classification and subscales were interesting, but not in the direction hypothesized. Patients in psychoanalysis were more likely than patients in psychodynamic psychotherapy to change from Secure to Preoccupied-entangled, to show a decrease in a positive representation of childhood (the loving scale) and a decrease in the overall positive state of mind and coherence with respect to attachment (positive state of mind scale). This is inconsistent with the theory that all psychotherapy causes a move towards security, particularly as most of the patients in psychoanalysis were symptomatic improvers. It contributes, however, to the theory proposed above that part of the process of an effective and intensive psychotherapeutic treatment in outpatients with personality disorders is to bring about more incoherence, perhaps akin to the psychoanalytic concepts of regression

and transference neurosis. On the other hand, there are several possible confounding factors or chance occurrences which could equally well explain this pattern. For example, the finding that the psychoanalysis group began with fewer Axis II subjects, lower anger, and higher positive state of mind may explain why these subjects were more likely to improve symptomatically and more likely to show a decrease during treatment in positive state of mind due to regression to the mean.

7.5.5 Overall improvement status

It is difficult to comment meaningfully on the association between initial AAI classification and improvement status due to the small sample size and the strong confound between improvement and treatment intensity. The lower rate of improvement in Dismissing subjects may be interpreted as confirming Horowitz's (1993) finding that such patients do less well in psychodynamic psychotherapy, but is more likely a result of the fact that all but one of these patients was in the less intensive treatment. An adequate study of the association between AAI classification and improvement would need to have sufficient numbers of subjects of each classification in each type of therapy. Undoubtedly, the initial AAI classification of the patient, the underlying psychopathology, and the parameters of the treatment (including intensity, process, and perhaps even the AAI classification of the therapist (Dozier et al., 1994)) interact in a complex way to influence treatment outcome.

The two significant findings on AAI change and improvement are small but interesting pieces of evidence in favor of the theory for improvement suggested above. Subjects who went from Dismissing to Secure during the course of a non-intensive treatment may have avoided the transition into preoccupation and incoherence that is a necessary component of a successful treatment, and leading to their failure to improve symptomatically. Despite the general trend towards greater insecurity in most of the subjects who improved, passivity did show a greater increase in non-improvers than

improvers, perhaps indicating that by termination one already has the suggestion from this scale that improvers are on their way to greater security.

7.5.6 Follow-along and follow-up analyses

The follow-along descriptive analysis, meaningfully limited to seven intensive improvers, satisfied one hypothesis and failed to confirm another. Intensive improvers, on average, did appear to grow significantly more insecure during treatment, as evidenced by a decrease in the positive SOM scale, and rebound towards termination. Despite this rebound, these subjects remained slightly (though not significantly) more insecure at termination than they had appeared at initial assessment. One conclusion from this may be that at termination, the incoherence of regression, transference neurosis, and relationship patterns being brought into the here and now, is such that even a subject who has improved significantly does not yet appear more secure than at initial assessment. Alternatively, it is possible that the apparent coherence at initial assessment is so high (due to the failure of the AAI as a clinical instrument) that it is unreasonable to expect coherence to rise above the initial level. The most useful evidence to address this question would be an adequate sample of follow-up data for subjects with follow-along data. Unfortunately, the seven subjects with follow-up data did not have follow-along data, and no significant differences were observed between their positive SOM scores by time point or by group. Larger sample sizes and more consistent data collection will be needed to address these questions.

7.5.7 AAI and YAWRS

The ultimate goal of process-outcome research, such as the study outlined in this thesis, is to move beyond individual associations between pairs of assessment, process, and outcome measures, and to create larger testable hypotheses about causal relationships in psychotherapy and psychoanalysis. It is therefore worthwhile, despite the small number of subjects with both AAI and YAWRS data, to look at the significant

results of the exploratory analysis and use it to build hypotheses that will be tested in future studies. The scattered associations between YAWRS factors and initial AAI scales may be accidental (as no Bonferroni correction was applied), but they may also suggest phenomena that are worthy of testing in a future study. Subjects with higher Pressure to Achieve were more likely to be rated by their analysts as producing good unconscious material and little resistance in the first year of treatment, possibly because they were compliant with the expectations of treatment. Subjects who reported higher anger in description of early relationships also were rated as producing good material and low resistance, perhaps because the analysts in this study, because of their theoretical stance, were particularly tuned to hearing about unconscious aggression. The negative association of AAI Neglecting, Coherence, and Positive State of Mind with dynamic material and technique is difficult to explain, particularly as there was no association with Reflective Function. If anything, we might have expected that the more coherent a narrative with which the patient begins, the more likely that dynamic material and technique would be high. Perhaps, the inclusive nature of the dynamic material and technique factors means that patients with incoherent material including sex and aggression scored higher on dynamic variables because more areas were covered on the YAWRS.

The most interesting finding concerning YAWRS and AAI is that, as expected, higher scores on dynamic treatment and technique predicted a change towards security (i.e., lower Passivity) by the termination AAI. This supports the initial hypothesis that the AAI may yield important information about structural change. Of course, the small sample size and the numerous possible confounders severely limits the generalizability of the finding and further research in this area is urgently needed. There is no ready explanation for why dynamic material and technique were not associated with Coherence or overall positive State of Mind scales. The unexpected association between supportive

interventions and a move towards security on both Passivity and overall State of Mind scales is exciting, as we had initially hypothesized (in Chapters 5 and 6) that supportive interventions are key to successful treatment (though the fact that they were not born out by the data earlier, led us to leave them out of the hypotheses in this chapter).

7.5.8 Overview of the model

Taken together, we propose a 12 step model for understanding the relationships between patient pathology, treatment intensity, treatment technique, attachment classification and subscales, and symptomatic improvement. The preceding 7 chapters have established a theoretical background for this model and provided tantalizing hints that parts are empirically supported in this quasi-experimental study.

- (1) Young adults with personality pathology, presenting with symptoms of depression and anxiety (Chapter 3)...
- (2) undergo psychoanalysis or psychodynamic psychotherapy. If...
- (3) dynamic technique (general, transference, and relationship interpretation) is begun in the first year of treatment, ...
- (4) supportive interventions are used to keep the alliance intact, and ...
- (5) this results in a high ratio of dynamic material (clear unconscious themes) to resistance (Chapter 6)...
- (6) patient classification on the AAI and SOM subscales move towards the “preoccupied/entangled” prototype during the course of treatment (Chapter 7)...
- (7) while symptom measures improve (Chapter 3).
- (8) As the patient progresses towards termination, symptoms continue to improve (Chapter 3)...
- (9) and AAI classification and SOM subscales indicate a move towards security (Chapter 7).
- (10) A year or more after termination, the patient continues to do well symptomatically

(Chapter 3)...

(11) and the AAI classification and SOM subscales show greater security than at initial assessment (Chapter 7).

(12) In a personality disordered population, this entire sequence is more likely to happen in psychoanalysis than in psychotherapy because (3) and (5) are more prevalent, resulting in more (6), (7), (8), (9), (10), and (11) (Chapter 5).

7.5.9 Study weaknesses

The most significant weaknesses of the AAI component of this study are the limitations of small sample size and failure of the AAI to adequately capture resolution with regards to trauma and loss. Even with the planned *n* of 25, the study had limited power, but this was compounded by missing assessments at termination, which further reduced the potential for exploring important patterns. As the numbers were small, it was necessary to use different overall samples for different statistical analyses, thus making it hard to compare and aggregate results from one analysis to another. Data for the follow-along and follow-up analyses were so scarce that only descriptive analyses were possible, and no subjects had both follow-along and follow-up assessments. Perhaps most seriously of all, small sample may have resulted in confounding factors (which may or may not have been associated with causal links) including the near total overlap between the psychoanalytic and symptomatic improver groups and the overrepresentation of Axis II, Preoccupied-entangled, and low positive SOM subjects in the psychodynamic psychotherapy group.

Given the interesting findings in the literature on the association of resolution with regards to trauma and loss and psychopathology (Dozier et al., 1999; Fonagy et al., 1996; Patrick et al., 1994), it was unfortunate that insufficient data were collected in the AAIs for this to be reliably coded in all assessments. More concerning still is the possibility that insufficient probing in this area led to a bias towards rating the patient Secure, which

could, in part, explain why a larger proportion than expected of patients at initial assessment fell into this category. This study would have benefited from redundant scoring of at least some interviews to verify the reliability of AAI classifications. This was impractical given that different coders were used for every different assessment, so as to prevent the coder from confusing information or patterns learned in one interview with those from another from the same patient. However, the fact that all of the AAI coders had established reliability through Main and Hesse's stringent requirements was felt to be a good indication of their reliability. Finally, the lack of existing psychometric data on the AAI subscales made analysis of these data highly speculative.

Given the importance of transference and regression in our understanding of AAI classification, it becomes important to be more consistent about the exact timing of initial and termination AAIs. Due to scheduling difficulties, as reported in Chapter 3, there was considerable variation in when these AAIs were performed: 5 of the 16 subjects with initial and termination AAIs had an initial AAI performed after the treatment had already begun (2 of which were done more than 5 months into treatment) in contrast with 11 subjects who had initial AAIs performed treatment began (including 3 who were seen 3 to 6 months before the first session). Meanwhile, 6 of 16 subjects had a termination AAI performed while they were still meeting with their treater (3 more than 7 months before the termination date) in contrast with 10 subjects seen after termination (including 5 interviewed between 4 and 6 months after the final session). While no clear patterns could be discerned in terms of assessment timing and AAI classification, greater care must be taken in standardising the administration of these interviews if the links between AAI classification, regression, transference, and structural change are to be clarified.

Data from the existing literature is still too sparse to make any conclusions specific to adolescent or young adult populations, such as were studied here. Thus questions

raised in Chapter 1 about the specific challenges and opportunities of patients in this age range, did not play a part in our interpretation of the data. Future studies would benefit from giving more attention to subject age.

7.5.10 Future research

As a small exploratory study, this research leaves much room for further work to verify its claims with larger samples and more careful data collection. Though it is expensive and time consuming, we believe that further research should again attempt to collect AAIs at initial, follow-along, termination, and follow-up in a large sample of subjects in psychotherapy and psychoanalysis. It is tempting to make use of less labor intensive measures of attachment, such as the Kobak Q-sort or one of several self-report measures (Crowell & Treboux, 1995; Stein, Jacobs, Ferguson, Allen, & Fonagy, 1998), and certainly information from studies using these measures for psychotherapy outcome would be interesting. However, evidence suggests that these measures are less effective at capturing stable underlying attachment structures and are more tied to superficial relationship patterns that would be less interesting in an assessment of structural change. If the AAI is applied in future studies it is crucial to ensure adequate probing according to the protocol, particularly of trauma and loss, so as to maximise the validity and consistency of the findings. Hopefully, future research will be able to directly address the question of how AAI classification and subscales change from initial assessment, through follow-along, termination, and follow-up, all in a large sample with a distribution of initial diagnoses, attachment classifications, and, treatment intensities.

Subscales of the AAI appear to be promising as continuous measures of probable experience and state of mind, and further psychometric data should be collected on these, in large samples, so as to have a better sense of normative values, reliability, and significance of change. Fonagy and Target's Reflective Function measure (1997) could also be usefully applied to the AAIs and help test Fonagy's hypothesis that reflective

function increases in successful psychotherapy (1995). Reflective function may represent a more central and theoretically meaningful construct than coherence, positive state of mind, or even security and thus, with empirical testing, could eventually be considered as a replacement for the AAI coding system. Unlike AAI security, RF may not show a decline during regression and transference because even as the patient gets more involved in their treatment and memories of the past, they are improving their capacity for mentalization. Thus RF may be more sensitive during therapy as a predictor of positive outcome, though potentially less useful as a measure of engagement and regression. Coding RF on the basis of a shortened interview could save some of the time and effort required for repeated application and coding of the AAI and make a long-term and large sample process-outcome study more feasible.

7.6 Conclusion

In this chapter, data from the AAI is presented as part of the Young Adult process-outcome study. Despite difficulties in consistently applying and coding the AAI at regular time intervals, the data collected show that it is a potentially useful measure for studying a patient's response to therapy before, during, and after termination. The high proportion of Secure ratings was surprising given previous research findings and is discussed in terms of the pathology of this sample and understanding how the AAI functions in a clinical sample. Analyses of the AAI with data from demographics and baseline assessment of psychopathology confirmed expectations from the literature (absence of association in the former, and strong associations in the latter). Most interestingly, patterns of change in the AAI suggested a new theory for the relationship between psychotherapy and state of mind with respect to attachment. We propose that in effective, intensive psychotherapy, particularly psychoanalysis, patients appear more preoccupied-entangled with respect to their childhood relationships (as reported in the

AAI) as part of a therapeutic process involving regression and transference neurosis.

Only at termination and beyond are these states of mind appropriately resolved and should then lead to an increased security, combined with symptomatic improvement. If, through further testing with larger and more complete samples, this theory is confirmed, the AAI, or related measures such as Fonagy's RF, will be indispensable for evaluating change in empirical studies of psychotherapy and psychoanalysis.

APPENDICES

Note: appendices are numbered to correspond to the chapters from which they are cited. At there are no appendices to Chapters 1 and 2, the section begins with Appendix 3.1

APPENDIX 3.1. TREATMENT PARAMETERS

<i>Subject</i>	<i>Possible Sessions</i>	<i>Sessions (%)</i>	<i>Analyst absence (%)</i>	<i>Patient cancellation (%)</i>	<i>Patient no-show (%)</i>	<i>Missing Data (%)</i>	<i>Estimated sessions attended</i>
A	1483	85	3	1	2	10	1401
B	290	75	4	21	0	0	218
C	1221	44	2	0	38	16	640
D	1129	71	10	1	10	8	871
E	1100	61	3	5	14	16	799
I	256	71	1	1	0	27	249
M	529	71	9	20	1	0	376
O	926	50	1	34	3	12	526
P	904	77	0	22	1	0	696
Q	679	66	5	18	0	10	498
T	1007	93	2	2	3	0	937
U	539	84	2	9	1	4	472
V	153	52	1	0	47	0	80
X	903	43	3	54	0	0	388

Table A3.1a-i. Attendance of intensive patients.

<i>Subject</i>	<i>Possible Sessions</i>	<i>Sessions (%)</i>	<i>Analyst absence (%)</i>	<i>Patient cancellation (%)</i>	<i>Patient no-show (%)</i>	<i>Missing Data (%)</i>	<i>Estimated sessions attended</i>
F	36	67	0	0	33	0	22
G	12	33	8	50	8	0	4
H	24	83	0	13	4	0	20
J	18	72	6	0	22	0	13
K	58	79	0	2	19	0	46
L	78	92	0	1	6	0	72
N	119	42	2	56	0	0	50
R	63	54	3	33	10	0	34
W	23	52	0	0	48	0	12
Y	83	76	0	23	1	0	63

Table A3.1a-ii. Attendance of non-intensive patients.

Test (df)	Gender (male)	Age	SES (1=highest class)	NART IQ
Tx intensity	.78 (1) n=25	1.1 (1,21)	-6.7* (1,19)	2.5 (1,12)
Tx length	3.1 (1,22)	-.09 (22)	-.41 (20)	.45 (14)
Time b/t 1st and last assessment	.11 (1,19)	.04 (20)	-.39 (18)	.32 (13)
# of assessments	5.4* (1,23)	.11 (23)	-.22 (21)	.61* (14)
% sessions attended	1.2 (1,22)	.43* (23)	-.15 (21)	-.12 (14)
# of sessions attended	2.7 (1,21)	.04 (22)	-.42 (20)	.35 (14)

Table A3.1b. Demographics versus treatment parameters.

Note: in the “Gender” column, a positive statistics indicates an association with maleness and in the “Tx intensity” row, a positive statistic indicates an association with intensive treatment

APPENDIX 3.2: RAW ASSESSMENT DATA

Legend: Initial = initial assessment, Follow n = follow-up assessment #n, Term = termination assessment, PT Follow n= post-termination follow-up assessment #n, ☺ = non-clinical range, ☹ = clinical range

SCL-90-R

<i>Subject</i>	<i>Assessments</i>					
A	INITIAL Missing	FOLLOW 1 2.25 ☹	FOLLOW 2 1.12 ☹	FOLLOW 3 1.13 ☹	TERM 0.27 ☺	PT FOLLOW 1 0.34 ☺
B	INITIAL Missing	TERM 0.02 ☺				
C	INITIAL Missing	FOLLOW 1 1.09 ☹	FOLLOW 2 0.69 ☹	FOLLOW 3 0.82 ☹	TERM 0.17 ☺	PT FOLLOW 1 0.27 ☺
D	INITIAL Missing	FOLLOW 1 1.45 ☹	FOLLOW 2 0.90 ☹	FOLLOW 3 1.30 ☹	TERM 2.19 ☹	PT FOLLOW 1 2.13 ☹
E	INITIAL Missing	FOLLOW 1 0.33 ☺	FOLLOW 2 0.31 ☺	FOLLOW 3 0.31 ☺	TERM 0.20 ☺	
M	INITIAL Missing	TERM 0.80 ☹	PT FOLLOW 1 0.26 ☺			
O	INITIAL Missing	FOLLOW 1 0.40 ☺	FOLLOW 2 0.07 ☺	FOLLOW 3 0.07 ☺	FOLLOW 4 0.08 ☺	TERM 0.04 ☺
P	INITIAL Missing	FOLLOW 1 0.94 ☹	FOLLOW 2 1.11 ☹	TERM 0.78 ☹	PT FOLLOW 1 2.11 ☹	
Q	INITIAL Missing	TERM 0.17 ☺	PT FOLLOW 1 0.12 ☺	PT FOLLOW 2 0.04 ☺	PT FOLLOW 3 0.06 ☺	
T	INITIAL Missing	FOLLOW 1 Missing	FOLLOW 2 1.57 ☹	TERM 0.89 ☹	PT FOLLOW 1 0.70 ☹	
U	INITIAL 1.72 ☹	FOLLOW 1 2.00 ☹	FOLLOW 2 1.94 ☹	TERM 1.38 ☹	PT FOLLOW 1 1.49 ☹	
V	INITIAL Missing	PT FOLLOW 1 0.58 ☺				
X	INITIAL Missing	FOLLOW 1 1.23 ☹	TERM 0.81 ☹			

Table A3.2a-i. SCL-90 by time for intensive patients.

<i>Subject</i>	<i>Assessments</i>			
F	INITIAL Missing	TERM 0.78 ☹	PT FOLLOW 1 1.40 ☹	PT FOLLOW 2 0.88 ☹
G	INITIAL 3.65 ☹			
H	INITIAL 2.72 ☹	TERM 0.95 ☹	PT FOLLOW 1 1.36 ☹	
J	INITIAL 0.57 ☺			
K	INITIAL 1.22 ☹	FOLLOW 1 1.82 ☹	FOLLOW 2 0.69 ☹	TERM 0.54 ☹
L	INITIAL 1.24 ☹	TERM 0.74 ☹	PT FOLLOW 1 1.17 ☹	
N	INITIAL Missing	TERM 2.33 ☹	PT FOLLOW 1 0.16 ☺	
R	INITIAL 1.58 ☹	TERM 1.34 ☹	PT FOLLOW 1 0.59 ☺	PT FOLLOW 2 0.28 ☺
S	INITIAL 0.54 ☺			
W	INITIAL 0.92 ☹			
Y	INITIAL 1.73 ☹	TERM 1.06 ☹		

Table 3.2a-ii. SCL-90 by time for non-intensive patients.

BDI

<i>Subject</i>	<i>Assessments</i>					
A	INITIAL 16 ☹️	FOLLOW 1 23 ☹️	FOLLOW 2 4 ☺️	FOLLOW 3 7 ☺️	TERM 0 ☺️	PT FOLLOW 1 Missing
B	INITIAL 0 ☺️	TERM 2 ☺️				
C	INITIAL 21 ☹️	FOLLOW 1 9 ☺️	FOLLOW 2 8 ☺️	FOLLOW 3 4 ☺️	TERM 3 ☺️	PT FOLLOW 1 1 ☺️
D	INITIAL 25 ☹️	FOLLOW 1 21 ☹️	FOLLOW 2 9 ☺️	FOLLOW 3 18 ☹️	TERM 34 ☹️	PT FOLLOW 1 16 ☺️
E	INITIAL 20 ☹️	FOLLOW 1 1 ☺️	FOLLOW 2 2 ☺️	FOLLOW 3 12 ☺️	TERM 1 ☺️	
I	INITIAL 29 ☹️					
M	INITIAL 29 ☹️	TERM 12 ☺️	PT FOLLOW 1 3 ☺️			
O	INITIAL 5 ☺️	FOLLOW 1 2 ☺️	FOLLOW 2 0 ☺️	FOLLOW 3 0 ☺️	FOLLOW 4 0 ☺️	TERM 0 ☺️
P	INITIAL 29 ☹️	FOLLOW 1 9 ☺️	FOLLOW 2 13 ☺️	TERM 9 ☺️	PT FOLLOW 1 20 ☹️	
Q	INITIAL 8 ☺️	TERM 0 ☺️	PT FOLLOW 1 0 ☺️	PT FOLLOW 2 0 ☺️	PT FOLLOW 3 0 ☺️	
T	INITIAL 25 ☹️	FOLLOW 1 39 ☹️	FOLLOW 2 36 ☹️	TERM 9 ☺️	PT FOLLOW 1 14 ☺️	
U	INITIAL 33 ☹️	FOLLOW 1 32 ☹️	FOLLOW 2 17 ☹️	TERM 15 ☺️	PT FOLLOW 1 13 ☺️	
V	INITIAL 27 ☹️	PT FOLLOW 1 9 ☺️				
X	INITIAL 12 ☺️	FOLLOW 1 18 ☹️	TERM 3 ☺️	PT FOLLOW 1 0 ☺️		

Table A3.2b-i. BDI by time for intensive patients.

<i>Subject</i>	<i>Assessments</i>			
F	INITIAL 25 ☹	TERM 10 ☺	PT FOLLOW 1 32 ☹	PT FOLLOW 2 20 ☹
G	INITIAL 43 ☹			
H	INITIAL 50 ☹	TERM 15 ☺	PT FOLLOW 1 21 ☹	
J	INITIAL 7 ☺			
K	INITIAL 9 ☺	FOLLOW 1 7 ☺	FOLLOW 2 0 ☺	TERM 4 ☺
L	INITIAL 14 ☺	TERM 19 ☹	PT FOLLOW 1 18 ☹	
N	INITIAL 15 ☺	TERM 20 ☹	PT FOLLOW 1 7 ☺	
R	INITIAL 31 ☹	TERM 17 ☹	PT FOLLOW 1 6 ☺	PT FOLLOW 2 7 ☺
S	INITIAL 8 ☺			
W	INITIAL 5 ☺			
Y	INITIAL 20 ☹	TERM 15 ☺		

Table A3.2b-ii. BDI by time for non-intensive patients.

STAI

Subject	Assessments					
	INITIAL	FOLLOW 1	FOLLOW 2	FOLLOW 3	TERM	PT FOLLOW 1
A	32 ☺	61 ☹	63 ☹	40 ☺	25 ☺	Missing
	59 ☹	49 ☹	57 ☹	45 ☹	38 ☺	Missing
B	INITIAL	TERM				
	33 ☺	37 ☺				
C	INITIAL	FOLLOW 1	FOLLOW 2	FOLLOW 3	TERM	PT FOLLOW 1
	49 ☹	39 ☺	34 ☺	34 ☺	28 ☺	24 ☺
D	INITIAL	FOLLOW 1	FOLLOW 2	FOLLOW 3	TERM	PT FOLLOW 1
	51 ☹	47 ☹	48 ☹	51 ☹	57 ☹	60 ☹
E	INITIAL	FOLLOW 1	FOLLOW 2	FOLLOW 3	TERM	
	47 ☹	35 ☺	36 ☺	44 ☹	36 ☺	
I	INITIAL					
	66 ☹					
M	INITIAL	TERM	PT FOLLOW 1			
	74 ☹	36 ☺	32 ☺			
O	INITIAL	FOLLOW 1	FOLLOW 2	FOLLOW 3	FOLLOW 4	TERM
	34 ☺	36 ☺	29 ☺	28 ☺	34 ☺	Missing
P	INITIAL	FOLLOW 1	FOLLOW 2	TERM	PT FOLLOW 1	
	57 ☹	47 ☹	45 ☹	45 ☹	70 ☹	
Q	INITIAL	TERM	PT FOLLOW 1	PT FOLLOW 2	PT FOLLOW 3	
	37 ☺	32 ☺	27 ☺	25 ☺	26 ☺	
T	INITIAL	FOLLOW 1	FOLLOW 2	TERM	PT FOLLOW 1	
	67 ☹	63 ☹	61 ☹	44 ☹	56 ☹	
U	INITIAL	FOLLOW 1	FOLLOW 2	TERM	PT FOLLOW 1	
	74 ☹	62 ☹	56 ☹	54 ☹	38 ☺	
V	INITIAL	PT FOLLOW 1				
	55 ☹	Missing				
X	INITIAL	FOLLOW 1	TERM	PT FOLLOW 1		
	47 ☹	62 ☹	39 ☺	28 ☺		

Table A3.2c-i. STAI by time for intensive patients. Upper = state, lower = trait.

<i>Subject</i>	<i>Assessments</i>			
	INITIAL	TERM	PT FOLLOW 1	PT FOLLOW 2
F	53 ☹	45 ☹	41 ☺	47 ☹
	60 ☹	45 ☹	52 ☹	49 ☹
G	INITIAL			
	71 ☹			
	74 ☹			
H	INITIAL	TERM	PT FOLLOW 1	
	66 ☹	66 ☹	48 ☹	
	72 ☹	56 ☹	54 ☹	
J	INITIAL			
	52 ☹			
	42 ☺			
K	INITIAL	FOLLOW 1	FOLLOW 2	TERM
	52 ☹	44 ☹	47 ☹	43 ☺
	55 ☹	49 ☹	46 ☹	52 ☹
L	INITIAL	TERM	PT FOLLOW 1	
	58 ☹	46 ☹	53 ☹	
	56 ☹	58 ☹	49 ☹	
N	INITIAL	TERM	PT FOLLOW 1	
	63 ☹	Missing	28 ☺	
	57 ☹	51 ☹	34 ☺	
R	INITIAL	TERM	PT FOLLOW 1	PT FOLLOW 2
	45 ☹	50 ☹	49 ☹	40 ☺
	68 ☹	61 ☹	56 ☹	47 ☹
S	INITIAL			
	47 ☹			
	47 ☹			
W	INITIAL			
	52 ☹			
	50 ☹			
Y	INITIAL	TERM		
	63 ☹	60 ☹		
	63 ☹	54 ☹		

Table A3.2c-ii. STAI by time for non-intensive patients. Upper = state, lower = trait.

SAS-M

<i>Subject</i>	<i>Assessments</i>					
	INITIAL Missing	FOLLOW 1	FOLLOW 2	FOLLOW 3	TERM	PT FOLLOW 1
A		2.58 ☹	2.00 ☹	1.87 ☺	1.61 ☺	1.64 ☺
B	INITIAL Missing	TERM 1.77 ☺				
C	INITIAL Missing	FOLLOW 1 1.45 ☺	FOLLOW 2 2.21 ☹	FOLLOW 3 1.96 ☹	TERM 1.23 ☺	PT FOLLOW 1 1.36 ☺
D	INITIAL Missing	FOLLOW 1 2.73 ☹	FOLLOW 2 2.14 ☹	FOLLOW 3 2.18 ☹	TERM 2.41 ☹	PT FOLLOW 1 2.70 ☹
E	INITIAL Missing	FOLLOW 1 1.70 ☺	FOLLOW 2 2.52 ☹	FOLLOW 3 1.93 ☹	TERM 1.76 ☺	
M	INITIAL Missing	TERM 1.87 ☺	PT FOLLOW 1 1.72 ☺			
O	INITIAL Missing	FOLLOW 1 1.84 ☺	FOLLOW 2 1.63 ☺	FOLLOW 3 1.58 ☺	FOLLOW 4 1.84 ☺	TERM 1.54 ☺
P	INITIAL Missing	FOLLOW 1 2.34 ☹	FOLLOW 2 2.50 ☹	TERM 2.36 ☹	PT FOLLOW 1 3.20 ☹	
Q	INITIAL Missing	TERM 1.68 ☺	PT FOLLOW 1 1.46 ☺	PT FOLLOW 2 1.32 ☺	PT FOLLOW 3 1.38 ☺	
T	INITIAL Missing	FOLLOW 1 2.64 ☹	FOLLOW 2 2.36 ☹	TERM 1.96 ☹	PT FOLLOW 1 2.00 ☹	
U	INITIAL 3.32 ☹	FOLLOW 1 2.89 ☹	FOLLOW 2 2.48 ☹	TERM 2.39 ☹	PT FOLLOW 1 1.85 ☺	
V	INITIAL Missing	PT FOLLOW 1 2.14 ☹				
X	INITIAL Missing	FOLLOW 1 2.00 ☹	TERM 1.92 ☹	PT FOLLOW 1 1.34 ☺		

Table A3.2d-i. SAS-M by time for intensive patients.

<i>Subject</i>	<i>Assessments</i>			
F	INITIAL Missing	TERM 2.14 ☹	PT FOLLOW 1 2.31 ☹	PT FOLLOW 2 2.56 ☹
G	INITIAL 3.50 ☹			
H	INITIAL 4.00 ☹	TERM 2.58 ☹	PT FOLLOW 1 2.14 ☹	
J	INITIAL 1.70 ☺			
K	INITIAL 2.36 ☹	FOLLOW 1 1.77 ☺	FOLLOW 2 2.03 ☹	TERM 2.11 ☹
L	INITIAL 3.06 ☹	TERM 3.00 ☹	PT FOLLOW 1 3.12 ☹	
N	INITIAL Missing	TERM 2.87 ☹	PT FOLLOW 1 1.62 ☺	
R	INITIAL 2.66 ☹	TERM 2.04 ☹	PT FOLLOW 1 2.13 ☹	PT FOLLOW 2 2.30 ☹
S	INITIAL 2.15 ☹			
W	INITIAL 2.03 ☹			
Y	INITIAL 2.48 ☹	TERM 2.55 ☹		

Table A3.2d-ii. SAS-M by time for non-intensive patients.

EPQ

<i>Subject</i>	<i>EPQ</i>				<i>BPRS</i>	<i>DIB</i>
	<i>E</i>	<i>P</i>	<i>N</i>	<i>L</i>	<i>Total</i>	<i>Total</i>
A	13 ☺	11 ☹	22 ☹	16	24	4 ☺
B	17 ☹	8 ☹	1 ☺	13	5	2 ☺
C	19 ☹	12 ☹	20 ☹	5	18	6 ☺
D	3 ☺	4 ☹	17 ☹	5	22	8 ☹
E	6 ☺	9 ☹	13 ☹	0	6	4 ☺
I	10 ☺	11 ☹	22 ☹	4	26	8 ☹
M	1 ☺	8 ☹	21 ☹	11	23	7 ☹
O	18 ☹	11 ☹	15 ☹	1		7 ☹
P	15 ☹	13 ☹	23 ☹	0	15	3 ☺
Q	10 ☺	8 ☹	16 ☹	6	17	5 ☺
T	0 ☺	4 ☹	18 ☹	9	25	4 ☺
U						2 ☺
V	7 ☺	11 ☹	19 ☹	3	20	6 ☺
X	16 ☹	10 ☹	23 ☹	5	10	4 ☺

Table A3.2e-i. EPQ, NART, BPRS, and DIB at initial assessment for intensive subjects.

<i>Subject</i>	<i>EPQ</i>				<i>BPRS</i>	<i>DIB</i>
	<i>E</i>	<i>P</i>	<i>N</i>	<i>L</i>	<i>Total</i>	<i>Total</i>
F	16 ☹	22 ☹	19 ☹	5	18	3 ☺
J					5	6 ☺
L						4 ☺
N	12 ☺	7 ☹	22 ☹	14	28	
R					2	
Y						3 ☺

Table A3.2e-ii. EPQ, NART, BPRS, and DIB at initial assessment for non-intensive subjects.

SADS-L

<i>Subject</i>	<i>Past</i>	<i>Assessments</i>					
		INITIAL	FOLLOW 1	FOLLOW 2	FOLLOW 3	TERM	PT FOLLOW 1
A	Maj dep Suicidal attempts	Maj dep Suicidal	Maj dep	None	None	None	None
B	None	INITIAL None	TERM Missing				
C	Hypomanic Drug abuse GAD Bulimia Violent	Cyclothymic Bulimia Drug abuse	Cyclothymic Bulimia	None	None	None	None
D	Maj dep Suicidal Drug abuse GAD	Maj dep Bulimia	Maj dep Bulimia	Maj dep	None	Maj dep Bulimia Suicidal	Maj dep Bulimia
E	Maj dep EtOH abuse Conduct disorder	Antisocial	None	None	Maj dep EtOH abuse Suicidal	EtOH abuse	
I	Maj dep Anorexia Violent	INITIAL Maj dep					
M	Maj dep Panic GAD Suicidal	Maj dep Panic GAD Suicidal	TERM None	PT FOLLOW 1 Maj dep Suicidal			
O	Maj dep Bipolar II EtOH abuse Violent	None	None	None	Maj dep	None	None
P	Maj dep	None	Maj dep	Maj dep	None	PT FOLLOW 1 Missing	
Q	Maj dep GAD	INITIAL None	TERM Missing	PT FOLLOW 1 GAD	PT FOLLOW 2 None	PT FOLLOW 3 None	
T	Maj dep GAD	Maj dep	Maj dep GAD Suicidal	Missing	None	PT FOLLOW 1 None	
U	Violent	INITIAL None	FOLLOW 1 Dysthymic	TERM Missing	PT FOLLOW 1 Missing		
V	Maj dep Suicide attempt Panic Bipolar II	Panic Bipolar II	PT FOLLOW 1 None				
X	Maj dep Panic GAD	Maj dep Panic GAD	Cyclothymic Panic	TERM None			

Table A3.2f-i. SADS-L by time for intensive patients.

<i>Subject</i>	<i>Past</i>	<i>Assessments</i>			
F	Dysthymic	INITIAL Dysthymic	TERM Dysthymic	PT FOLLOW 1 Dysthymic	PT FOLLOW 2 GAD
G	EtOH abuse Suicide attempts	INITIAL Dysthymic OCD			
H	Dythymic Anorexia Suicide attempt	INITIAL Dysthymic	TERM None	PT FOLLOW 1 Missing	
J	Dysthymic Suicide attempt	INITIAL None			
K	Social phobia	INITIAL None	FOLLOW 1 None	FOLLOW 2 None	TERM None
L	Dysthymic	INITIAL Dysthymic	TERM None	PT FOLLOW 1 Missing	
N	Maj dep Suicidal Panic OCD	INITIAL Maj dep Suicidal Panic OCD	TERM Dysthymic	PT FOLLOW 1 None	
R	Dysthymic Bulimia	INITIAL Dysthymic Bulimia	TERM Dysthymic Bulimia	PT FOLLOW 1 Anorexia	PT FOLLOW 2 Missing
S	Dysthymic	INITIAL Dysthymic			
W	Dysthymic Anorexia Bulimia	INITIAL Dysthymic			
Y	Dysthymic Suicidal	INITIAL Dysthymic	TERM Dysthymic		

Table A3.2f-ii. SADS-L by time for non-intensive patients.

SCID-II

<i>Subject</i>	<i>Assessments</i>					
A	INITIAL Self-defeating Narcissistic Borderline	FOLLOW 1 Self-defeating Paranoid	FOLLOW 2 Narcissistic	FOLLOW 3 Paranoid Conduct	TERM Paranoid	PT FOLLOW 1 None
B	INITIAL None	TERM None				
C	INITIAL Borderline	FOLLOW 1 None	FOLLOW 2 Antisocial	FOLLOW 3 Narcissistic Borderline Antisocial	TERM None	PT FOLLOW 1 None
D	INITIAL None	FOLLOW 1 Avoidant Borderline	FOLLOW 2 None	FOLLOW 3 None	TERM None	PT FOLLOW 1 Self-defeating
E	INITIAL None	FOLLOW 1 Conduct	FOLLOW 2 None	FOLLOW 3 Borderline Conduct	TERM None	
I	INITIAL Dependent Borderline					
M	INITIAL None	TERM Avoidant	PT FOLLOW 1 Avoidant			
O	INITIAL Borderline	FOLLOW 1 None	FOLLOW 2 None	FOLLOW 3 None	FOLLOW 4 None	TERM Antisocial PD NOS
P	INITIAL Dependent	FOLLOW 1 None	FOLLOW 2 None	TERM Self-defeating	PT FOLLOW 1 Passive- aggressive Paranoid Narcissistic Borderline	
Q	INITIAL Passive- aggressive	TERM None	PT FOLLOW 1 None	PT FOLLOW 2 None	PT FOLLOW 3 None	
T	INITIAL Avoidant Schizotypal	FOLLOW 1 Avoidant	FOLLOW 2 Avoidant	TERM None	PT FOLLOW 1 None	
U	INITIAL NOS	FOLLOW 1 Narcissistic	FOLLOW 2 Missing	TERM Missing	PT FOLLOW 1 Missing	
V	INITIAL Conduct	PT FOLLOW 1 None				
X	INITIAL None	FOLLOW 1 Passive- aggressive Histrionic Borderline	TERM None	PT FOLLOW 1 Missing		

Table A3.2g-i. SCID-II by time for intensive patients.

<i>Subject</i>	<i>Assessments</i>			
F	INITIAL Self-defeating	TERM Self-defeating	PT FOLLOW 1 Self-defeating	PT FOLLOW 2 Missing
G	INITIAL Self-defeating Borderline Conduct			
H	INITIAL Self-defeating	TERM None	PT FOLLOW 1 Missing	
K	INITIAL Avoidant	FOLLOW 1 Avoidant	FOLLOW 2 Avoidant	TERM Avoidant
L	INITIAL Avoidant	TERM Avoidant	PT FOLLOW 1 Missing	
N	INITIAL Avoidant Obsessive- compulsive Self-defeating Paranoid	TERM None	PT FOLLOW 1 None	
R	INITIAL Avoidant	TERM Avoidant	PT FOLLOW 1 Missing	PT FOLLOW 2 Missing
S	INITIAL Paranoid			
W	INITIAL Paranoid Borderline			
Y	INITIAL Avoidant	TERM Avoidant		

Table A3.2g-ii. SCID-II by time for non-intensive patients.

APPENDIX 3.3. ASSESSMENT MEASURES

Test (df)	Gender (Male)	Age	SES (1=highest class)	NART IQ
BDI	-.74 (1,23)	.20 (23)	.43 (21)	.38 (14)
STAI-S	-3.1 (1,23)	.07 (23)	.42 (21)	.17 (14)
STAI-T	-.80 (1,23)	.02 (23)	.34 (21)	.42 (14)
# Ax I Diags	-6.6* (1,23)	-.25 (23)	.29 (21)	.22 (14)
# Ax II Diags	-.20 (1,23)	-.17 (21)	.36 (19)	-.05 (14)
Mood	-10.0** (1) n=25	-.50 (1,21)	1.7 (1,19)	.08 (1,12)
Clus A	-2.6 (1) n=23	-.29 (1,19)	1.5 (1,17)	-.05 (1,12)
Clus B	.29 (1) n=23	.00 (1,19)	-.95 (1,17)	.11 (1,12)
Clus C	-.02 (1) n=23	.58 (1,19)	1.4 (1,17)	.00 (1,12)
BPRS	-.85 (1,14)	-.24 (16)	.29 (15)	-.06 (13)
DIB	-5.7* (1,16)	-.18 (17)	.21 (16)	-.09 (14)
EPQ-E	3.2 (1,13)	.17 (15)	-.20 (14)	-.02 (14)
EPQ-P	3.8 (1,13)	.26 (15)	-.14 (14)	.35 (14)
EPQ-N	-2.3 (1,13)	-.21 (15)	.20 (14)	.40 (14)
EPQ-L	-.15 (1,13)	-.25 (15)	.28 (14)	-.36 (14)

Table A3.3a. Demographics versus initial assessments

Note: in the “gender” column, a positive statistic indicates an association between “maleness” and the variable in the corresponding row

Test (df)	Tx intensity	Tx length	Time b/ t 1* & last assessment	# of assessments	% sessions attended	# of sessions attended
BDI	-.02 (1,23)	-.15 (24)	-.23 (21)	.10 (25)	.05 (24)	-.05 (23)
STAI-S	-.98 (1,23)	-.33 (24)	-.34 (21)	-.25 (25)	.06 (24)	-.29 (23)
STAI-T	-.27 (1,23)	-.08 (24)	-.17 (21)	.10 (25)	.09 (24)	-.00 (23)
# Ax I Diags	.03 (1,23)	.15 (24)	.25 (21)	.07 (25)	-.43* (24)	.03 (23)
# Ax II Diags	-2.8 (1,21)	-.05 (22)	.19 (20)	-.03 (23)	-.07 (22)	-.05 (21)
Mood	-.78 (1) n=25	-.54 (1,22)	.06 (1,19)	-.82 (1,23)	-.81 (1,22)	-.29 (1,21)
Clus A	-1.1 (1) n=23	-.72 (1,20)	.28 (1,18)	-.51 (1,21)	-.81 (1,20)	-.30 (1,19)
Clus B	.05 (1) n=23	-.73 (1,20)	2.7 (1,18)	-.05 (1,21)	-.27 (1,20)	.71 (1,19)
Clus C	-2.3 (1) n=23	-1.5 (1,20)	-.83 (1,18)	-.02 (1,20)	1.7 (1,20)	-1.3 (1,19)
BPRS	.78 (1,14)	.22 (16)	.28 (14)	.16 (16)	.24 (16)	.25 (16)
DIB	.83 (1,16)	.06 (17)	.30 (16)	-.09 (18)	-.06 (18)	.02 (17)
EPQ-E	-.56 (1,13)	.05 (15)	-.04 (14)	-.04 (15)	-.46 (15)	-.20 (15)
EPQ-P	-3.0 (1,13)	-.27 (15)	-.31 (14)	-.11 (15)	-.22 (15)	-.34 (15)
EPQ-N	-.42 (1,13)	.11 (15)	.31 (14)	.10 (15)	-.12 (15)	.03 (15)
EPQ-L	-.82 (1,13)	-.08 (15)	.06 (14)	-.15 (15)	.16 (15)	.03 (15)

Table A3.3b. Treatment parameters versus initial assessments

Note: in the “Tx intensity” column, a positive statistic indicates an association between psychoanalysis and the variable in the corresponding row

Test (df)	BDI	STAI-S	STAI-T	# Ax I Diags	# Ax II Diags	Mood	Clus A	Clus B	Clus C	BPRS	DIB	EPQ-E	EPQ-P	EPQ-N
STAI-S	.71** (44)													
STAI-T	.86** (45)	.83** (44)												
# Ax I Diags	.35* (43)	.33* (42)	.40** (43)											
# Ax II Diags	.21 (42)	.27 (41)	.33* (42)	.27 (41)										
Mood	4.8* (1,42)	2.4 (1,41)	4.7* (1,42)	24.2** (1,41)	2.1 (1,4)									
Clus A	.72 (1,40)	1.1 (1,39)	.02 (1,40)	.51 (1,39)	12.5** (1,40)	.51 (1) n=42								
Clus B	.01 (1,40)	.01 (1,39)	.10 (1,40)	.05 (1,39)	6.8* (1,40)	.00 (1) n=42	.59 (1) n=42							
Clus C	4.5* (1,40)	3.3 (1,39)	6.5* (1,40)	.11 (1,39)	12.7** (1,40)	.91 (1) n=42	.00 (1) n=42	.12 (1) n=42						
BPRS	.35 (16)	.51* (16)	.50* (16)	.41 (16)	.66* (16)	6.7* (1,14)	3.5 (1,12)	1.4 (1,12)	4.1 (1,12)					
DIB	.09 (18)	.04 (18)	.12 (18)	.37 (18)	-.06 (17)	1.3 (1,13)	.13 (1,15)	.88 (1,15)	.81 (1,15)	.48 (14)				
EPQ-E	-.51 (15)	-.57* (15)	-.55* (15)	-.21 (15)	.10 (15)	1.0 (1,13)	1.4 (1,13)	2.7 (1,13)	.00 (1,13)	-.38 (14)	-.34 (14)			
EPQ-P	.15 (15)	-.11 (15)	.05 (15)	-.17 (15)	-.03 (15)	.01 (1,13)	2.8 (1,13)	.50 (1,13)	.60 (1,13)	-.14 (14)	-.29 (14)	.58* (15)		
EPQ-N	.61* (15)	.50 (15)	.75** (15)	.50 (15)	.40 (15)	6.7* (1,13)	.26 (1,13)	.47 (1,13)	2.2 (1,13)	.62* (14)	.33 (14)	-.10 (15)	.20 (15)	
EPQ-L	-.26 (15)	-.02 (15)	-.04 (15)	.41 (15)	.50 (15)	1.9 (1,13)	2.5 (1,13)	.00 (1,13)	.78 (1,13)	.35 (14)	-.22 (14)	-.08 (15)	-.28 (15)	-.09 (15)

Table A3.3c. Initial + termination assessment inter-relationships

APPENDIX 3.4. ASSESSMENT DATA CHANGE SCORES

Legend: RC = statistically reliable change, CC = clinically significant change
 imp = improvement, det = deterioration, nc = no change
 clin = remains in clinical range, non-clin = remains in non-clinical range
 (# Imp-#Det) / init clin = # of scales showing improvement (RC + CC) minus number
 of scales showing deterioration (RC+CC) divided by # of scales in the clinical range at
 initial assessment

<i>Subject</i>	<i>BDI</i>	<i>STAI-T</i>	<i># Axis I</i>	<i># Axis II</i>	<i>(#Imp-#Det) Init clin</i>	<i>Status</i>
A	RC imp CC non-clin	RC imp CC imp	RC imp CC imp	RC imp CC clin	6 / 3	Improver
B	RC nc CC non-clin	RC nc CC non-clin	Missing	RC nc CC non-clin	0 / 0	Non-improver
C	RC imp CC imp	RC imp CC imp	RC imp CC imp	RC nc CC imp	7 / 4	Improver
D	RC nc CC clin	RC nc CC clin	RC nc CC clin	RC nc CC non-clin	0 / 3	Non-improver
E	RC imp CC imp	RC nc CC imp	RC nc CC clin	RC nc CC non-clin	3 / 3	Improver
M	RC imp CC imp	RC imp CC imp	RC imp CC imp	RC nc CC det	6 / 3	Improver
O	RC nc CC non-clin	RC imp CC imp	RC nc CC non-clin	RC nc CC clin	2 / 2	Improver
P	RC imp CC imp	RC imp CC clin	RC nc CC non-clin	RC nc CC clin	3 / 3	Improver
Q	RC nc CC non-clin	RC nc CC imp	RC nc CC non-clin	RC nc CC imp	2 / 2	Improver
T	RD imp CC imp	Missing	RC nc CC imp	RC imp CC imp	5 / 3	Improver
U	RC imp CC imp	RC nc CC clin	Missing	Missing	2 / 2	Improver
X	RC imp CC non-clin	RC nc CC imp	RC imp CC imp	RC nc CC non-clin	4 / 2	Improver

Table A3.4a-i. Termination change by scales and overall status for intensive patients.

<i>Subject</i>	<i>BDI</i>	<i>STAI-T</i>	<i># Axis I</i>	<i># Axis II</i>	<i>(#Imp-#Det) Init clin</i>	<i>Status</i>
F	RC imp CC imp	RC nc CC clin	RC nc CC clin	RC nc CC clin	2 / 4	Non-improver
H	RC imp CC imp	RC imp CC clin	RC nc CC imp	RC nc CC imp	5 / 4	Improver
K	RC nc CC non-clin	RC nc CC clin	RC nc CC non-clin	RC nc CC clin*	0 / 2	Non-improver
L	RC nc CC det	RC nc CC clin	RC nc CC imp	RC nc CC clin	0 / 2	Non-improver
N	RC nc CC det	RC nc CC clin	RC imp CC clin	RC imp CC imp	3 / 3	Improver
R	RC imp CC clin	RC nc CC clin	RC nc CC clin	RC nc CC clin*	1 / 4	Non-improver
Y	RC nc CC imp	RC nc CC clin	RC nc CC clin	RC nc CC clin	1 / 4	Non-improver

* = missing data substituted on # Axis II using SCI-II self-report questionnaire

Table A3.4a-ii. Termination change by scales and overall status for non-intensive patients.

<i>Subject</i>	<i>BDI</i>	<i>STAI-T</i>	<i># Axis I</i>	<i># Axis II</i>	<i>(#Imp-#Det) Init clin</i>	<i>Status</i>
A	Missing	Missing	RC imp CC imp	RC imp CC imp	4 / 2	Improver sustained
C	RC imp CC imp	RC imp CC imp	RC imp CC imp	RC nc CC imp	7 / 4	Improver sustained
D	RC nc CC imp	RC nc CC clin	RC nc CC clin	RC nc CC non-det	0 / 3	Non-improver sustained
M	RC imp CC imp	RC imp CC imp	RC imp CC clin	RC nc CC det	5 / 3	Improver sustained
P	RC nc CC clin	RC nc CC clin	Missing	RC det CC clin	0 / 3	Non-improver deteriorated
Q	RC nc CC non-clin	RC imp CC imp	RC nc CC det	RC nc CC imp	2 / 2	Improver sustained
T	RD nc CC imp	RC imp CC clin	RC nc CC imp	RC imp CC imp	5 / 4	Improver sustained
U	RC imp CC imp	RC imp CC clin	Missing	Missing	3 / 2	Improver sustained

Table A3.4b-i. Follow-up change by scales and overall status for intensive patients.

<i>Subject</i>	<i>BDI</i>	<i>STAI-T</i>	<i># Axis I</i>	<i># Axis II</i>	<i>(#Imp-#Det)</i> <i>Init clin</i>	<i>Status</i>
F	RC nc CC clin	RC nc CC clin	RC nc CC clin	RC nc CC clin	0 / 4	Non-improver sustained
H	RC imp CC clin	RC imp CC clin	Missing	Missing	2 / 2	Improver sustained
L	RC nc CC det	RC nc CC clin	Missing	Missing	0 / 1	Non-improver sustained
N	RC nc CC non-clin	RC imp CC imp	RC imp CC imp	RC imp CC imp	6 / 3	Improver sustained
R	RC imp CC imp	RC imp CC clin	RC nc CC clin	Missing	3 / 3	Improver improved

Table 3.4b-ii. Follow-up change by scales and overall status for non-intensive patients.

*** = missing data substituted on # Axis II using SCI-II self-report questionnaire**

APPENDIX 3.5. VARIABLES ASSOCIATED WITH OVERALL IMPROVEMENT STATUS.

	<i>Improvers (mean, sd)</i>	<i>Non-imp (mean, sd)</i>	<i>Statistic</i>
Demographics			
Gender (male)	5 male 7 female	3 male 4 female	$\chi^2 = 0.003$ (df=1, n=19)
Age	22.8 (1.4)	22.7 (2.4)	t = -0.14 (df=17)
SES (1=highest class)	2.2 (1.6)	2.0 (1.0)	t = -0.25 (df=17)
NART IQ	121.6 (2.1)	121.1 (5.2)	t = -0.28 (df=10)
Treatment parameters			
Tx intensity	10 intensive 2 non-intens	2 intensive 5 non-intens	$\chi^2 = 5.7$ (1, n=19)*
Tx length (weeks)	258 (126)	163 (139)	t = -0.15 (df=17)
# of assessments	4.8 (1.0)	3.7 (1.4)	t = -1.8 (df = 17)
% sessions attended	73 (19)	76 (12)	t = 0.47 (df = 17)
# of sessions attended	567 (376)	189 (308)	t = -2.2 (df = 17)*

* = p < 0.05

	<i>Improvers (mean, sd)</i>	<i>Non-imp (mean, sd)</i>	<i>Statistic</i>
Baseline assessments			
BDI	22 (12)	18 (11)	t = -0.75 (df = 17)
STAI-T	58 (10)	57 (12)	t = -0.34 (df=17)
# Ax I Diags	1.6 (1.6)	1.0 (0.8)	t = -0.91 (df=17)
# Ax II Diags	1.3 (1.2)	0.67 (0.51)	t = -1.1 (df=16)
Mood disorder	11 yes, 1 no	5 yes, 2 no	$\chi^2 = 1.4$ (1, n=19)
Suicidality	4 yes, 8 no	2 yes, 5 no	$\chi^2 = 0.05$ (1, n=19)
Substance abuse	3 yes, 9 no	1 yes, 6 no	$\chi^2 = 0.31$ (1, n=19)
Eating disorder	2 yes, 10 no	2 yes, 5 no	$\chi^2 = 0.38$ (1, n=19)
Anxiety disorder	6 yes, 6 no	2 yes, 5 no	$\chi^2 = 1.5$ (1, n=19)
BPRS	18 (7)	12 (10)	t = -1.4 (df=11)
DIB	4.6 (1.6)	4.0 (2.3)	t = -0.58 (df=13)
EPQ-E	11 (7)	12 (8)	t = 0.22 (df=11)
EPQ-P	9 (3)	11 (9)	t = 0.66 (df=11)
EPQ-N	19 (4)	12 (10)	t = -2.0 (df=11)
EPQ-L	7 (6)	8 (5)	t = 0.27 (df=11)

* = p < 0.05

APPENDIX 4.1. THE YOUNG ADULT WEEKLY RATING SCALE

The Anna Freud Centre Young Adult Weekly Rating Scale

Case initials: _____ Weekly number: _____ Rater's Initials: _____ Week ending: _____

Please circle every day that a session took place.

If the session was cancelled, enter 'A' for analyst's cancellation or 'P' for patient's cancellation.

Mon: _____ Tues: _____ Weds: _____ Thurs: _____ Fri: _____

Try and rate *all* items. If you feel that an item is inapplicable, please circle round all points 0 1 2 3.

I. GENERAL CHARACTERISTICS

A. RESISTANCE

Choose statement which is most consistent with patient's current stance.

- | | |
|--|-----|
| 1. Patient on the whole is relatively ready to explore thoughts and feelings. | Y N |
| 2. Free association shows strong signs of inhibition, lacking in spontaneity. | Y N |
| 3. Patient is actively resisting examining thoughts, feelings and motivations related to problems. | Y N |
| 4. Unexpected or unexplained intensification of resistance. | Y N |

If YES to (3) or (4), indicate specific nature of resistance:

	<u>PRES</u>	<u>INT</u>
5. Patient is silent.	0 1 2 3	0 1 2 3
6. Patient is consciously withholding selected material.	0 1 2 3	0 1 2 3
7. Lies to the analyst (distorts reality deliberately).	0 1 2 3	0 1 2 3
8. Brings mental or actual lists of issues to discuss into the session.	0 1 2 3	0 1 2 3
9. Determined to retain focus on the conscious meaning of his associations.	0 1 2 3	0 1 2 3
10. Speaks of many things without exploring any in depth.	0 1 2 3	0 1 2 3
11. Material is repetitive, predictable and boring.	0 1 2 3	0 1 2 3
12. Is intolerant of analyst's version of reality.	0 1 2 3	0 1 2 3
13. Doesn't listen to the analyst's interpretations.	0 1 2 3	0 1 2 3
14. Other (please specify).	0 1 2 3	0 1 2 3

B. GENERAL ATTITUDE TO ANALYSIS

1. Is enjoying the analysis.	0 1 2 3	0 1 2 3
2. Seems dissatisfied with analysis.	0 1 2 3	0 1 2 3
3. Wants to withdraw from (end) the analysis.	0 1 2 3	0 1 2 3
4. Feels in conflict about external commitments and commitment to analysis.	0 1 2 3	0 1 2 3
5. Patient attacks or abuses the analyst.	0 1 2 3	0 1 2 3
6. Patient is in analytic impasse.	0 1 2 3	0 1 2 3

C. TIME KEEPING.

Presence (PRES): 0 = absent
 1 = possibly present
 2 = present
 3 = prominent

Interpretation (INT): 0 = not taken up
 1 = taken into account in other interpretation
 2 = taken up directly
 3 = a central interpretation of the week

1. Patient keeps to the normal times. Y N

If NO, then:

2. Patient is significantly early (at least 10 minutes). 0 1 2 3 0 1 2 3
 3. Late (at least 3-4 minutes). 0 1 2 3 0 1 2 3
 4. Has difficulty leaving. 0 1 2 3 0 1 2 3
 5. Leaves during session. 0 1 2 3 0 1 2 3
 6. Ends the session (but roughly on time). 0 1 2 3 0 1 2 3

D. MISSED SESSIONS

1. Patient misses session or sessions. Y N

If YES, then:

2. Was there notice of missed sessions? Y N
 3. Was there reality reason given? Y N

If YES to (3) then, indicate the nature of the conflict:

4. Analyst's judgement of validity of reason (Y=valid). Y N
 a. Work 0 1 2 3 0 1 2 3
 b. Relationship - sexual. 0 1 2 3 0 1 2 3
 c. Relationship - non-sexual. 0 1 2 3 0 1 2 3
 d. Reaction to threat within analysis. (If YES and clear, please specify nature of threat): 0 1 2 3 0 1 2 3

5. Attempt to establish specific pattern of relationship with analyst through absence. 0 1 2 3 0 1 2 3
 6. Possible meaning of absence taken up by patient (1st column) and by analyst (2nd column). 0 1 2 3 0 1 2 3
 7. Other irregularities of attendance (please specify): 0 1 2 3 0 1 2 3

E. PATIENT'S BEHAVIOUR IN SESSION

1. Patient keeps to analytic parameters. Y N

If NO, then:

2. Patient excessively fidgets or moves on the couch. 0 1 2 3 0 1 2 3
 3. Patient is immobile and hardly moves. 0 1 2 3 0 1 2 3
 4. Is drowsy and incoherent. 0 1 2 3 0 1 2 3
 5. Makes demand for physical contact (e.g. handshake, holding). 0 1 2 3 0 1 2 3
 6. Jumps up from couch in anxiety or anger. 0 1 2 3 0 1 2 3
 7. Brings excessive quantity of physical objects into sessions (books, suitcases, bags). 0 1 2 3 0 1 2 3
 8. Falls asleep during session. 0 1 2 3 0 1 2 3
 9. Does not use couch. 0 1 2 3 0 1 2 3
 10. Lies on couch facing analyst. 0 1 2 3 0 1 2 3
 11. Patient enacts during the session. Y N

Presence (PRES): 0 = absent
 1 = possibly present
 2 = present
 3 = prominent

Interpretation (INT): 0 = not taken up
 1 = taken into account in other interpretation
 2 = taken up directly
 3 = a central interpretation of the week

If YES, please specify:

F. QUALITY OF ANALYTIC MATERIAL

1. Patient is capable of bringing helpful analytic material. Y N

If YES, then:

- | | | |
|---|---------|---------|
| 2. There is a clear theme to the material. | 0 1 2 3 | 0 1 2 3 |
| 3. Patient reflects on content of his free association. | 0 1 2 3 | 0 1 2 3 |
| 4. Makes links between different themes within session or between sessions. | 0 1 2 3 | 0 1 2 3 |
| 5. Reports dreams. | 0 1 2 3 | 0 1 2 3 |
| 6. Reports daydreams and fantasies. | 0 1 2 3 | 0 1 2 3 |
| 7. Produces associations to dreams and fantasies. | 0 1 2 3 | 0 1 2 3 |
| 8. Thinks over meaning or significance of his/her thoughts and fantasies. | 0 1 2 3 | 0 1 2 3 |

G. PATIENT'S AGGRESSION AND SEXUALITY IN THE ANALYSIS

1. Patient's aggression and sexuality in the analysis goes beyond domain of fantasy. Y N

If YES, then:

- | | | |
|---|---------|---------|
| 2. Patient plays with body products during session (picks ears, nose, nails). | 0 1 2 3 | 0 1 2 3 |
| 3. Resorts to auto-erotic gratification (thumb-sucking, rocking, masturbation). | 0 1 2 3 | 0 1 2 3 |
| 4. Behaves highly sexually seductively to the analyst. | 0 1 2 3 | 0 1 2 3 |
| 5. Intrudes into analyst's life (follows analyst, contacts relations). | 0 1 2 3 | 0 1 2 3 |
| 6. Shouts at the analyst. | 0 1 2 3 | 0 1 2 3 |
| 7. Physically threatens analyst. | 0 1 2 3 | 0 1 2 3 |
| 8. Patient is frightened that he/she will feel like physically threatening the analyst. | 0 1 2 3 | 0 1 2 3 |
| 9. Enacts his/her aggressive impulses on furniture etc. during session. | 0 1 2 3 | 0 1 2 3 |
| 10. Destructive of premises where analyst works (steals, shouts, smears, attacks other patients). | 0 1 2 3 | 0 1 2 3 |

H. MATURITY OF PATIENT'S MENTAL FUNCTIONING

- | | |
|--|-----|
| 1. Patient's mental capacity is adequate to the analytic task. | Y N |
| 2. Quality of mental functioning fluctuates between mature and immature modes. | Y N |
| 3. Mental functioning shifts in response to interpretation. | Y N |
| 4. Mental functions mostly immature and causes difficulties for the analysis. | Y N |

If YES to (2), (3), or (4), then specify:

- | | | |
|---|---------|---------|
| 5. Patient's material is dominated by confusion, disorder. | 0 1 2 3 | 0 1 2 3 |
| 6. Primitive defences predominate over mature ones (splitting, projection, projective identification, denial vs. repression, inhibition, displacement, reaction formation, humour). | 0 1 2 3 | 0 1 2 3 |

Presence (PRES): 0 = absent
 1 = possibly present
 2 = present
 3 = prominent

Interpretation (INT): 0 = not taken up
 1 = taken into account in other interpretation
 2 = taken up directly
 3 = a central interpretation of the week

- | | | |
|--|---------|---------|
| 7. Capacity to differentiate his/her own intentions and wishes from that of his/her objects appears lacking. | 0 1 2 3 | 0 1 2 3 |
| 8. Has difficulty in understanding metaphoric meaning and becomes concrete. | 0 1 2 3 | 0 1 2 3 |
| 9. Material does not have an 'as if' quality (e.g. fantasy feels too real). | 0 1 2 3 | 0 1 2 3 |
| 10. Patient 'regresses' to childlike form of thinking in session. | 0 1 2 3 | 0 1 2 3 |
| 11. Patient's personal boundaries appear fluid. | 0 1 2 3 | 0 1 2 3 |
| 12. Not able to reflect on his/her own mental state. | 0 1 2 3 | 0 1 2 3 |

II. MANIFEST CONTENT

A. BODY

Choose *one* of the following statements:

- | | |
|---|-----|
| 1. Bodily state and functions totally absent from the material. | Y N |
| 2. Mentioned occasionally. | Y N |
| 3. Constitute a significant proportion of manifest material. | Y N |

If YES to (3) (or (2) if thought significant), please complete 'A2' and 'A3'

A2. CHARACTERISED BY:

- | | |
|---|-----|
| 4. Generally satisfaction with body appearance or function. | Y N |
| 5. Neutral attitude. | Y N |
| 6. Dislike, disgust or dissatisfaction. | Y N |
| 7. Extreme pride/excitement. | Y N |
| 8. Other (please specify): | Y N |
-

A3. SPECIFIC TOPICS

- | | | |
|---|---------|---------|
| 9. Physical symptoms of health (e.g. headache, sleep disturbances). | 0 1 2 3 | 0 1 2 3 |
| 10. Actual bodily damage. Please specify area: | 0 1 2 3 | 0 1 2 3 |
| <hr/> | | |
| 11. Physical appearance or adequacy of body parts (e.g. beauty, ugliness etc.). | 0 1 2 3 | 0 1 2 3 |
| 12. Excretory body functions (urination, defecation). | 0 1 2 3 | 0 1 2 3 |
| 13. Eating, drinking, diet etc. | 0 1 2 3 | 0 1 2 3 |
| 14. Fantasies about damage to analyst's body. | 0 1 2 3 | 0 1 2 3 |
| 15. Fantasies about possibility of damage to any part of his/her body. | 0 1 2 3 | 0 1 2 3 |
| 16. Confusion about parts of the body. | 0 1 2 3 | 0 1 2 3 |
| 17. Other (please specify). | 0 1 2 3 | 0 1 2 3 |
-

B. SELF-ESTEEM

- | | |
|---|-----|
| 1. Self is seen as on the whole adequate. | Y N |
|---|-----|

If NO, please indicate

- | | | |
|--|---------|---------|
| 2. Negative perceptions of self by others is dominant theme. | 0 1 2 3 | 0 1 2 3 |
|--|---------|---------|

Presence (PRES): 0 = absent
1 = possibly present
2 = present
3 = prominent

Interpretation (INT): 0 = not taken up
1 = taken into account in other interpretation
2 = taken up directly
3 = a central interpretation of the week

- | | | |
|---|---------|---------|
| 3. Positive perceptions of self by others is dominant theme. | 0 1 2 3 | 0 1 2 3 |
| 4. Patient on the whole feels inadequate and inferior. | 0 1 2 3 | 0 1 2 3 |
| 5. Patient on the whole feels particularly effective and superior. | 0 1 2 3 | 0 1 2 3 |
| 6. Patient's self-esteem seems to fluctuate in extreme way. | 0 1 2 3 | 0 1 2 3 |
| 7. There is marked grandiosity (e.g. patient boasts or exaggerates achievements). | 0 1 2 3 | 0 1 2 3 |
| 8. Patient is acutely self-conscious. | 0 1 2 3 | 0 1 2 3 |

C. HISTORICAL MATERIAL

- | | | |
|--|---|---|
| 1. Historical material (childhood, adolescence) is a part of the material. | Y | N |
|--|---|---|

If YES, please indicate if material is:

- | | | |
|--|---------|---------|
| 2. a. In the main new (fresh). | | |
| b. A new perspective on memories already recalled. | 0 1 2 3 | 0 1 2 3 |
| c. Repetition of material from previous sessions. | 0 1 2 3 | 0 1 2 3 |
| d. Please note memory of abuse. | 0 1 2 3 | 0 1 2 3 |
| 3. Also indicate (your best guess) if memories are of: | | |
| a. Infancy (0-2). | 0 1 2 3 | 0 1 2 3 |
| b. Early childhood (2-4). | 0 1 2 3 | 0 1 2 3 |
| c. Infant school (5-8). | 0 1 2 3 | 0 1 2 3 |
| d. Junior school (8-11). | 0 1 2 3 | 0 1 2 3 |
| e. Pre-adolescence (12-14). | 0 1 2 3 | 0 1 2 3 |
| f. Adolescence (15-18). | 0 1 2 3 | 0 1 2 3 |
| g. Young adulthood. | 0 1 2 3 | 0 1 2 3 |
| 4. And if affective tone of memories is generally: | | |
| a. Happy/joyous. | 0 1 2 3 | 0 1 2 3 |
| b. Sad/depressive. | 0 1 2 3 | 0 1 2 3 |
| c. Anxious/guilty. | 0 1 2 3 | 0 1 2 3 |
| d. Terrifying/traumatic. | 0 1 2 3 | 0 1 2 3 |
| e. Angry | 0 1 2 3 | 0 1 2 3 |
| f. Confused. | 0 1 2 3 | 0 1 2 3 |
| g. Excited | 0 1 2 3 | 0 1 2 3 |
| h. Resentful | 0 1 2 3 | 0 1 2 3 |
| i. Terror | 0 1 2 3 | 0 1 2 3 |
| j. Disgust | 0 1 2 3 | 0 1 2 3 |
| k. Affect is absent. | 0 1 2 3 | 0 1 2 3 |

D. RELATIONSHIP THEMES WITHIN FAMILY

- | | | |
|---|---|---|
| 1. Patient talks about relationship themes with his/her family. | Y | N |
|---|---|---|

If YES, then:

- | | | |
|--|---------|---------|
| 2. Discussion concerns relationships in the: | | |
| a. Past | 0 1 2 3 | 0 1 2 3 |
| b. Present | 0 1 2 3 | 0 1 2 3 |

Presence (PRES): 0 = absent
1 = possibly present
2 = present
3 = prominent

Interpretation (INT): 0 = not taken up
1 = taken into account in other interpretation
2 = taken up directly
3 = a central interpretation of the week

3. With:
- | | | |
|----------------------------|---------|---------|
| a. Mother | 0 1 2 3 | 0 1 2 3 |
| b. Father | 0 1 2 3 | 0 1 2 3 |
| c. Sibling(s) | 0 1 2 3 | 0 1 2 3 |
| d. Other (please specify): | 0 1 2 3 | 0 1 2 3 |
-

4. Please specify relationship theme:
- | | | |
|--|---------|---------|
| a. Being loved, admired or accepted. | 0 1 2 3 | 0 1 2 3 |
| b. Being punished or threatened. | 0 1 2 3 | 0 1 2 3 |
| c. Being unloved, neglected or rejected. | 0 1 2 3 | 0 1 2 3 |
| d. Inhibition on closeness. | 0 1 2 3 | 0 1 2 3 |
| e. Being separate and independent. | 0 1 2 3 | 0 1 2 3 |
| f. Fearing loss or loss of love. | 0 1 2 3 | 0 1 2 3 |
| g. Being sole object of object's affect. | 0 1 2 3 | 0 1 2 3 |
| h. Being in a conflict of loyalties. | 0 1 2 3 | 0 1 2 3 |
| i. Being rivalrous with object. | 0 1 2 3 | 0 1 2 3 |
| j. Identification with positive aspects. | 0 1 2 3 | 0 1 2 3 |
| k. Identification with negative aspects. | 0 1 2 3 | 0 1 2 3 |
| l. Fear of identification with family. | 0 1 2 3 | 0 1 2 3 |
| m. Envious relation with object. | 0 1 2 3 | 0 1 2 3 |
| n. Being found out. | 0 1 2 3 | 0 1 2 3 |
| o. Being the object of abuse. | 0 1 2 3 | 0 1 2 3 |
| p. Being the abuser. | 0 1 2 3 | 0 1 2 3 |
| q. Other (please specify). | 0 1 2 3 | 0 1 2 3 |

E. RELATIONSHIP THEMES WITH FRIENDS

1. Patient discusses relationship themes with his/her friends. Y N

If YES:

2. Discussion concerns relationship in the:
- | | | |
|------------|---------|---------|
| a. Past. | 0 1 2 3 | 0 1 2 3 |
| b. Present | 0 1 2 3 | 0 1 2 3 |

3. With:
- | | | |
|----------------------------|---------|---------|
| a. Same sex friend | 0 1 2 3 | 0 1 2 3 |
| b. Opposite sex friend | 0 1 2 3 | 0 1 2 3 |
| c. Peer Group | 0 1 2 3 | 0 1 2 3 |
| d. Other (please specify): | 0 1 2 3 | 0 1 2 3 |
-

4. Please specify relationship theme:
- | | | |
|--|---------|---------|
| a. Being threatened or attacked. | 0 1 2 3 | 0 1 2 3 |
| b. Being loved, admired or accepted. | 0 1 2 3 | 0 1 2 3 |
| c. Being punished or threatened. | 0 1 2 3 | 0 1 2 3 |
| d. Being unloved, neglected or rejected. | 0 1 2 3 | 0 1 2 3 |
| e. Inhibition on closeness. | 0 1 2 3 | 0 1 2 3 |

Presence (PRES): 0 = absent
 1 = possibly present
 2 = present
 3 = prominent

Interpretation (INT): 0 = not taken up
 1 = taken into account in other interpretation
 2 = taken up directly
 3 = a central interpretation of the week

f. Being separate and independent.	0 1 2 3	0 1 2 3
g. Fear of being separate and independent.	0 1 2 3	0 1 2 3
h. Fearing loss or loss of love.	0 1 2 3	0 1 2 3
i. Being sole object of object's affect.	0 1 2 3	0 1 2 3
j. Being in a conflict of loyalties.	0 1 2 3	0 1 2 3
k. Being rivalrous with object.	0 1 2 3	0 1 2 3
l. Belonging in (being accepted by) a group.	0 1 2 3	0 1 2 3
m. Fear of isolation.	0 1 2 3	0 1 2 3
n. Fear of falling behind peer group and its values.	0 1 2 3	0 1 2 3
o. Other (please specify):	0 1 2 3	0 1 2 3

F. SEXUAL RELATIONS

1. Discusses relationship themes with sexual partner. Y N

If YES, then:

2. Discussion concerns relationship in the:

a. Past	0 1 2 3	0 1 2 3
b. Present	0 1 2 3	0 1 2 3

3. With:

a. Same sex	0 1 2 3	0 1 2 3
b. Opposite sex	0 1 2 3	0 1 2 3

4. Please indicate relationship themes:

a. Being threatened or attacked.	0 1 2 3	0 1 2 3
b. Being loved, admired or accepted.	0 1 2 3	0 1 2 3
c. Inhibition on closeness, fear of intimacy.	0 1 2 3	0 1 2 3
d. Being punished or threatened.	0 1 2 3	0 1 2 3
e. Being unloved, neglected or rejected.	0 1 2 3	0 1 2 3
f. Inhibition on closeness.	0 1 2 3	0 1 2 3
g. Being separate and independent.	0 1 2 3	0 1 2 3
h. Fearing loss or loss of love.	0 1 2 3	0 1 2 3
i. Being sole object of object's affect.	0 1 2 3	0 1 2 3
j. Being in a conflict of loyalties.	0 1 2 3	0 1 2 3
k. Being rivalrous with object.	0 1 2 3	0 1 2 3
l. Wish to form stable couple and planning future together.	0 1 2 3	0 1 2 3
m. Thoughts of having children.	0 1 2 3	0 1 2 3
n. Fear of being controlled, overwhelmed, trapped.	0 1 2 3	0 1 2 3
o. Other (please specify):	0 1 2 3	0 1 2 3

G. MONEY AND WORK

1. Patient talks about plans and difficulties with money and work matters. Y N

If YES, then:

2. Problems with managing their money.	0 1 2 3	0 1 2 3
3. Fear of losing employment.	0 1 2 3	0 1 2 3

Presence (PRES): 0 = absent
1 = possibly present
2 = present
3 = prominent

Interpretation (INT): 0 = not taken up
1 = taken into account in other interpretation
2 = taken up directly
3 = a central interpretation of the week

- | | | |
|---|---------|---------|
| 4. Anxiety about gaining employment. | 0 1 2 3 | 0 1 2 3 |
| 5. Preconception about being in right job. | 0 1 2 3 | 0 1 2 3 |
| 6. Discussion of career plans. | 0 1 2 3 | 0 1 2 3 |
| 7. Problems with owing money. | 0 1 2 3 | 0 1 2 3 |
| 8. Conflicts with financial support from parents. | 0 1 2 3 | 0 1 2 3 |

H. CURRENT LIFE EVENTS

- | | | |
|--|---------|------------|
| 1. Patient's current or recent life situation is significant part of patient's material during any part of the week (e.g. successes and failures, initiation or breakdown of relationship, finance, education, employment, loss, pregnancy, health, drugs, legal or other problems). | | Y N |
| If YES, then: | | |
| 2. Nature of event or situation. Please specify: | 0 1 2 3 | 0 1 2 3 |
| _____ | | |
| _____ | | |
| 3. Seriousness of event/situation: | | |
| a. Extremely serious (life threatening). | 0 1 2 3 | 0 1 2 3 |
| b. Very serious (long term consequences). | 0 1 2 3 | 0 1 2 3 |
| c. Quite serious (threatening but limited). | 0 1 2 3 | 0 1 2 3 |
| d. No threat to life but psychologically threatening. | 0 1 2 3 | 0 1 2 3 |
| e. No threat. | 0 1 2 3 | 0 1 2 3 |
| 4. Patient's capacity to cope with event: | | |
| a. Realistically discusses options (coping). | 0 1 2 3 | 0 1 2 3 |
| b. Unrealistic in appraisal of situation - exaggerated seriousness and difficulty. | 0 1 2 3 | 0 1 2 3 |
| c. Ineffective coping strategies. | 0 1 2 3 | 0 1 2 3 |
| d. Other (please specify): | 0 1 2 3 | 0 1 2 3 |
| _____ | | |
| 5. Patient's affective reaction to event/situation. | | |
| a. Happy/joyous | 0 1 2 3 | 0 1 2 3 |
| b. Frightened (anxious), panicky (catastrophising) | 0 1 2 3 | 0 1 2 3 |
| c. Depressed (hopeless) | 0 1 2 3 | 0 1 2 3 |
| d. Suicidal | 0 1 2 3 | 0 1 2 3 |
| e. Angry/raging. | 0 1 2 3 | 0 1 2 3 |
| f. Lacking in affect (neutral). | 0 1 2 3 | 0 1 2 3 |

I. GENDER, AGE, AND RACE ISSUES

- | | | |
|---|---------|------------|
| 1. Did the material contain reference to gender, age or race of patient or analyst? | | Y N |
| If YES, then: | | |
| 2. Indicate whether material concerns: | 0 1 2 3 | 0 1 2 3 |
| a. Analyst's gender | 0 1 2 3 | 0 1 2 3 |
| b. Patient's gender | 0 1 2 3 | 0 1 2 3 |
| c. Analyst's age | 0 1 2 3 | 0 1 2 3 |
| d. Analyst's cultural origin | 0 1 2 3 | 0 1 2 3 |

Presence (PRES): 0 = absent
1 = possibly present
2 = present
3 = prominent

Interpretation (INT): 0 = not taken up
1 = taken into account in other interpretation
2 = taken up directly
3 = a central interpretation of the week

- | | | |
|------------------------------|---------|---------|
| e. Patient's cultural origin | 0 1 2 3 | 0 1 2 3 |
| f. Analyst's colour | 0 1 2 3 | 0 1 2 3 |
| g. Patient's colour | | |

J. ADULT IDENTITY

- | | | |
|--|---|---|
| 1. In the analyst's view, a new adult concern appears, or an aspect of adult identity is strengthened. | Y | N |
|--|---|---|

If YES, please indicate content:

2. Content:

- | | | |
|---|---------|---------|
| a. Planning future work activities | 0 1 2 3 | 0 1 2 3 |
| b. Thinking about starting a family | 0 1 2 3 | 0 1 2 3 |
| c. Shift in sexual identity | 0 1 2 3 | 0 1 2 3 |
| d. Distancing from dependence on parents | 0 1 2 3 | 0 1 2 3 |
| e. Increased sense of self-sufficiency and personal independence. | 0 1 2 3 | 0 1 2 3 |

3. Please indicate your view of this change:

- | | | |
|---|---------|---------|
| a. Mature shift | 0 1 2 3 | 0 1 2 3 |
| b. Defensive manoeuvre | 0 1 2 3 | 0 1 2 3 |
| c. An early phase of positive change, but achieved through a primitive defensive manoeuvre (e.g. identification with analyst, splitting, etc.). | 0 1 2 3 | 0 1 2 3 |
| d. Other (please specify): | 0 1 2 3 | 0 1 2 3 |
-

K. SEXUALITY

- | | | |
|---|---|---|
| 1. The subject of sexuality is deliberately avoided by the patient. | Y | N |
| 2. Sexuality is not referred to in patient's material. | Y | N |
| 3. Overt material concerning sexuality of self and/or other is significant theme. | Y | N |
| 4. Analyst thinks that sexual concerns are implied in material but not explicit. | Y | N |

If YES to (3) or (4):

5. Indicate if material concerns:

- | | | |
|--|---------|---------|
| a. Sexual fantasy about analyst (overtly expressed). | 0 1 2 3 | 0 1 2 3 |
| b. Having sexual awareness. | 0 1 2 3 | 0 1 2 3 |
| c. Memories of childhood sexuality. | 0 1 2 3 | 0 1 2 3 |
| d. Genital sexuality (intercourse). | 0 1 2 3 | 0 1 2 3 |
| e. Preenatal sexuality. | 0 1 2 3 | 0 1 2 3 |
| f. Masturbation | 0 1 2 3 | 0 1 2 3 |
| g. Promiscuity | 0 1 2 3 | 0 1 2 3 |
| h. Homosexuality | 0 1 2 3 | 0 1 2 3 |
| i. Birth/pregnancy | 0 1 2 3 | 0 1 2 3 |
| j. Perverse enactments: | | |
| i) Masochistic | 0 1 2 3 | 0 1 2 3 |
| ii) Sadistic | 0 1 2 3 | 0 1 2 3 |
| iii) Fetishistic | 0 1 2 3 | 0 1 2 3 |

Presence (PRES): 0 = absent
1 = possibly present
2 = present
3 = prominent

Interpretation (INT): 0 = not taken up
1 = taken into account in other interpretation
2 = taken up directly
3 = a central interpretation of the week

- | | | |
|---|---------|---------|
| iv) Anal/urinary | 0 1 2 3 | 0 1 2 3 |
| v) Scopophilia/exhibitionism. | 0 1 2 3 | 0 1 2 3 |
| k. Enactments in the analytic session (e.g. looks in closed box on analyst's bookcase). | 0 1 2 3 | 0 1 2 3 |

6. Material in the analyst's judgement motivated by:

- | | | |
|---|---------|---------|
| a. Inhibition about sexuality. | 0 1 2 3 | 0 1 2 3 |
| b. Wish fulfilling fantasy. | 0 1 2 3 | 0 1 2 3 |
| c. Anxiety about failure or difficulties. | 0 1 2 3 | 0 1 2 3 |
| d. Shame about wishes/fantasies. | 0 1 2 3 | 0 1 2 3 |
| e. Ambivalence about sexual identity. | 0 1 2 3 | 0 1 2 3 |
| f. Realistic account of experiences. | 0 1 2 3 | 0 1 2 3 |

L. DISCUSSION OF TREATMENT PARAMETERS

- | | | |
|--|---------|---------|
| 1. a. Patient discusses thoughts/fantasies surrounding treatment parameters (scheduling treatment times, changes, holiday breaks, missed sessions etc.). | 0 1 2 3 | 0 1 2 3 |
| b. There is <i>no</i> discussion of treatment parameters when this would be expected. | 0 1 2 3 | 0 1 2 3 |

If YES to (a) (i.e. rating 1, 2 or 3 in the 'Presence' column), please specify:

2. Parameter concerned is:

- | | | |
|----------------------------|---------|---------|
| a. Ending of sessions. | 0 1 2 3 | 0 1 2 3 |
| b. Times of sessions. | 0 1 2 3 | 0 1 2 3 |
| c. Weekend. | 0 1 2 3 | 0 1 2 3 |
| d. Holiday break. | 0 1 2 3 | 0 1 2 3 |
| e. Termination. | 0 1 2 3 | 0 1 2 3 |
| f. Other (please specify): | 0 1 2 3 | 0 1 2 3 |
-

3. Associated affect:

- | | | |
|----------------------------|---------|---------|
| a. Happy | 0 1 2 3 | 0 1 2 3 |
| b. Sad | 0 1 2 3 | 0 1 2 3 |
| c. Angry | 0 1 2 3 | 0 1 2 3 |
| d. Resentful | 0 1 2 3 | 0 1 2 3 |
| e. Guilty | 0 1 2 3 | 0 1 2 3 |
| f. Anxious | 0 1 2 3 | 0 1 2 3 |
| g. Confused | 0 1 2 3 | 0 1 2 3 |
| h. Notably absent | 0 1 2 3 | 0 1 2 3 |
| i. Other (please specify): | 0 1 2 3 | 0 1 2 3 |
-

III. PRECONSCIOUS THEMES

(These refer to attitudes which may be conscious or preconscious to the patient, i.e. can be made conscious through simple interpretation).

A. GENERAL TRANSFERENCE THEMES

- | | | |
|---|---------|---------|
| 1. a. Analytic relationship is central conscious theme. | 0 1 2 3 | 0 1 2 3 |
|---|---------|---------|

Presence (PRES): 0 = absent
1 = possibly present
2 = present
3 = prominent

Interpretation (INT): 0 = not taken up
1 = taken into account in other interpretation
2 = taken up directly
3 = a central interpretation of the week

- | | | | |
|---|--|---------|---------|
| b. | Patient reluctant to talk about relationship with analyst. | 0 1 2 3 | 0 1 2 3 |
| c. | Analyst notes and interprets nature of patient's feelings about transferences. | 0 1 2 3 | 0 1 2 3 |
| 2. Positive wishes towards analyst (approach): | | Y N | |
| If YES, then: | | | |
| a. | Affectionate and loving feelings. | 0 1 2 3 | 0 1 2 3 |
| b. | Intense curiosity about analyst. | 0 1 2 3 | 0 1 2 3 |
| c. | Wish for greater intimacy and closeness. | 0 1 2 3 | 0 1 2 3 |
| d. | Wish to be analyst's sole object (also jealousy). | 0 1 2 3 | 0 1 2 3 |
| e. | Wish to identify with analyst. | 0 1 2 3 | 0 1 2 3 |
| f. | Wish for dependence on analyst. | 0 1 2 3 | 0 1 2 3 |
| g. | Patient idealises analyst. | 0 1 2 3 | 0 1 2 3 |
| h. | Erotic feelings about analyst. | 0 1 2 3 | 0 1 2 3 |
| 3. Transference with anxiety (avoidance): | | Y N | |
| If YES, then: | | | |
| a. | Wary and suspicious of analyst. | 0 1 2 3 | 0 1 2 3 |
| b. | Fears rejection (loss of analyst's regard). | 0 1 2 3 | 0 1 2 3 |
| c. | Fears attack from analyst. | 0 1 2 3 | 0 1 2 3 |
| d. | Fears being manipulated by analyst. | 0 1 2 3 | 0 1 2 3 |
| e. | Fears seduction. | 0 1 2 3 | 0 1 2 3 |
| f. | Fears abandonment. | 0 1 2 3 | 0 1 2 3 |
| g. | Fears own rage and anger towards analyst. | 0 1 2 3 | 0 1 2 3 |
| h. | Fears criticism. | 0 1 2 3 | 0 1 2 3 |
| i. | Fears intimacy and intrusion. | 0 1 2 3 | 0 1 2 3 |
| 4. Transferences with competitive and aggressive themes (fighting): | | Y N | |
| If YES, then: | | | |
| a. | Competitive and rivalrous quality of analytic experience. | 0 1 2 3 | 0 1 2 3 |
| b. | Triumphant and victorious over analyst or analysis. | 0 1 2 3 | 0 1 2 3 |
| c. | Feeling dominated by analyst. | 0 1 2 3 | 0 1 2 3 |
| d. | Patient overtly expresses feelings of aversion towards analyst. | 0 1 2 3 | 0 1 2 3 |
| e. | Patient overtly expresses fear of aggression from analyst. | 0 1 2 3 | 0 1 2 3 |
| f. | Feeling analyst is helpless and vulnerable. | 0 1 2 3 | 0 1 2 3 |
| g. | Feeling analyst has been taken over. | 0 1 2 3 | 0 1 2 3 |
| 5. Transferences with resentment (injury): | | Y N | |
| If YES, then: | | | |
| a. | Wish for total care and sense of deprivation. | 0 1 2 3 | 0 1 2 3 |
| b. | Not being understood. | 0 1 2 3 | 0 1 2 3 |
| c. | Having been abandoned, deserted or rejected. | 0 1 2 3 | 0 1 2 3 |
| d. | Hating analyst. | 0 1 2 3 | 0 1 2 3 |
| e. | Derogatory/denigrating of analyst. | 0 1 2 3 | 0 1 2 3 |

Presence (PRES): 0 = absent
 1 = possibly present
 2 = present
 3 = prominent

Interpretation (INT): 0 = not taken up
 1 = taken into account in other interpretation
 2 = taken up directly
 3 = a central interpretation of the week

- | | | |
|---|---------|---------|
| f. Undermining and attacking of analyst. | 0 1 2 3 | 0 1 2 3 |
| g. Expresses regret and wishes to undo damage inflicted on analyst. | 0 1 2 3 | 0 1 2 3 |
| h. Complains about demands of analyst. | 0 1 2 3 | 0 1 2 3 |
| i. Not being special for the analyst. | 0 1 2 3 | 0 1 2 3 |

6. Primitive transferences:

Y N

If YES, then:

- | | | |
|---|---------|---------|
| a. Patient feels part of analyst (merging). | 0 1 2 3 | 0 1 2 3 |
| b. Analyst not felt to be human. | 0 1 2 3 | 0 1 2 3 |
| c. Totally dependent on analyst. | 0 1 2 3 | 0 1 2 3 |
| d. Patient feels in control of analyst (omnipotence). | 0 1 2 3 | 0 1 2 3 |
| e. Patient feels totally controlled by analyst. | 0 1 2 3 | 0 1 2 3 |
| f. Feels let down, betrayed and deceived by analyst. | 0 1 2 3 | 0 1 2 3 |

B. CHANGES OF TRANSFERENCE ACROSS THE WEEK

(This refers to conscious or preconscious content)

1. Transference more or less consistent across the week.

Y N

If NO, then:

- | | | |
|---|---------|---------|
| a. Rapidly shifts within the session. | 0 1 2 3 | 0 1 2 3 |
| b. Shifts between sessions. | 0 1 2 3 | 0 1 2 3 |
| c. Shifts in response to interpretation. | 0 1 2 3 | 0 1 2 3 |
| d. Gradually becomes clearer over the week. | 0 1 2 3 | 0 1 2 3 |

C. PREDOMINANT PRECONSCIOUS EMOTIONAL STANCE

(This refers to conscious or preconscious content)

1. a. Lack of affect in patient's material.

Y N

b. Clear emotional stance in material.

Y N

If YES to (b), indicate nature of affect (Questions 2 - 4):

2. Sadness (Depressive position)

Y N

If YES, then:

- | | | |
|---|---------|---------|
| a. Depression concerning genuine mourning (lost opportunity, loss of love). | 0 1 2 3 | 0 1 2 3 |
| b. Sadness concerned with damage to the object. | 0 1 2 3 | 0 1 2 3 |
| c. Guilt concerning sexual wishes. | 0 1 2 3 | 0 1 2 3 |
| d. Guilt concerning aggressive wishes and actions. | 0 1 2 3 | 0 1 2 3 |
| e. Realistic pleasure | 0 1 2 3 | 0 1 2 3 |

3. Primitive emotional stance (including paranoid schizoid position).

Y N

If YES, then:

- | | | |
|---|---------|---------|
| a. Depression based on persecutory internal object. | 0 1 2 3 | 0 1 2 3 |
| b. Fear of punishment. | 0 1 2 3 | 0 1 2 3 |
| c. Experience of being attacked. | 0 1 2 3 | 0 1 2 3 |
| d. Manic defence | 0 1 2 3 | 0 1 2 3 |
| e. Omnipotence | 0 1 2 3 | 0 1 2 3 |

Presence (PRES): 0 = absent
1 = possibly present
2 = present
3 = prominent

Interpretation (INT): 0 = not taken up
1 = taken into account in other interpretation
2 = taken up directly
3 = a central interpretation of the week

- | | | |
|--|---------|---------|
| f. Striking absence of guilt. | 0 1 2 3 | 0 1 2 3 |
| g. Fear of annihilation (destruction of self). | 0 1 2 3 | 0 1 2 3 |
| h. Fear of disintegration (including fear of being mad). | 0 1 2 3 | 0 1 2 3 |
| i. Fear of being humiliated. | 0 1 2 3 | 0 1 2 3 |
| j. Intense shame | 0 1 2 3 | 0 1 2 3 |
| k. Fear of loss of love. | 0 1 2 3 | 0 1 2 3 |

4. Anger

Y N

If YES, then:

- | | | |
|---|---------|---------|
| a. Inappropriate anger towards outside figures. | 0 1 2 3 | 0 1 2 3 |
| b. Anger towards analyst. | 0 1 2 3 | 0 1 2 3 |
| c. Striking absence of anger. | 0 1 2 3 | 0 1 2 3 |
| d. Narcissistic rage. | 0 1 2 3 | 0 1 2 3 |

D. PREDOMINANT DEFENCES

(These are manifest to analyst but preconscious to patient)

- | | | |
|---|---------|---------|
| 1. Distant from his/her feelings (isolation). | 0 1 2 3 | 0 1 2 3 |
| 2. Attributes difficulties to external influences (externalisation). | 0 1 2 3 | 0 1 2 3 |
| 3. Experiences own feelings as coming from the object (projection). | 0 1 2 3 | 0 1 2 3 |
| 4. Expresses/enacts wishes appropriate to a younger person, (child or infant) (regression). | 0 1 2 3 | 0 1 2 3 |
| 5. Reverses feelings/thoughts into their opposite (reaction formation). | 0 1 2 3 | 0 1 2 3 |
| 6. Gives excessive rational explanations (intellectualisation). | 0 1 2 3 | 0 1 2 3 |
| 7. Uses playfulness in order to avoid awareness of painful affect (humour). | 0 1 2 3 | 0 1 2 3 |
| 8. Fragments, destroys links, breaks up meaning (form of splitting). | 0 1 2 3 | 0 1 2 3 |
| 9. Experiences part of self as different (other) person (form of splitting). | 0 1 2 3 | 0 1 2 3 |
| 10. Self alternates between distinct identities without awareness. | 0 1 2 3 | 0 1 2 3 |
| 11. Denies significance of idea (denial). | 0 1 2 3 | 0 1 2 3 |
| 12. Isolation of affect. | 0 1 2 3 | 0 1 2 3 |
| 13. Reverses an impulse into its opposite (e.g. turning passive into active). | 0 1 2 3 | 0 1 2 3 |
| 14. Feelings experienced by the analyst caused by the patient (counter transference/projective identification). | 0 1 2 3 | 0 1 2 3 |
| 15. Other (please specify): | 0 1 2 3 | 0 1 2 3 |
-

IV. PREDOMINANT UCS DYNAMIC THEMES

(These are UCS themes which the analyst is aware of but the patient is not. If the patient acknowledges these, he/she does so only after interpretation and working through.)

A. GENERAL

1. Are there clearly discernable UCS themes?

Y N

Presence (PRES): 0 = absent
1 = possibly present
2 = present
3 = prominent

Interpretation (INT): 0 = not taken up
1 = taken into account in other interpretation
2 = taken up directly
3 = a central interpretation of the week

2. Are these related to aspects of object relations? Y N

If YES to (1), please indicate content:

- | | | |
|--|---------|---------|
| 3. About loss of object. | 0 1 2 3 | 0 1 2 3 |
| 4. Destructive greed or envy of the object. | 0 1 2 3 | 0 1 2 3 |
| 5. Being the sole object of mother's affection (mother/infant couple). | 0 1 2 3 | 0 1 2 3 |
| 6. Wish to possess and control the object which is deeply feared. | 0 1 2 3 | 0 1 2 3 |
| 7. Wish to possess and control the object which provides safety, admiration, love and nourishment. | 0 1 2 3 | 0 1 2 3 |
| 8. Be separate and independent from the object. | 0 1 2 3 | 0 1 2 3 |
| 9. Demand to be admired by object. | 0 1 2 3 | 0 1 2 3 |

B. ARE THESE REACTIONS TO AGGRESSION? Y N

If YES, please indicate content.

- | | | |
|---|---------|---------|
| 1. Immediate reality reasons for the patient being angry. | 0 1 2 3 | 0 1 2 3 |
| 2. Sadism towards the the object. | 0 1 2 3 | 0 1 2 3 |
| 3. Establishing dominance and control over object. | 0 1 2 3 | 0 1 2 3 |
| 4. Spoiling and envious aggression. | 0 1 2 3 | 0 1 2 3 |
| 5. Aggression as defence against anxiety. | 0 1 2 3 | 0 1 2 3 |
| 6. Aggression to defend the integrity of the self. | 0 1 2 3 | 0 1 2 3 |
| 7. Aggression as self-assertion. | 0 1 2 3 | 0 1 2 3 |
| 8. Aggression against mental functioning of analyst. | 0 1 2 3 | 0 1 2 3 |
| 9. Wish to obliterate the object (e.g. omnipotence). | 0 1 2 3 | 0 1 2 3 |
| 10. Aggression turned against self. | 0 1 2 3 | 0 1 2 3 |
| 11. Contamination of sexuality with aggression. | 0 1 2 3 | 0 1 2 3 |
| 12. Identifies with aggressor and being the victim of aggression. | 0 1 2 3 | 0 1 2 3 |

C. ARE THESE RELATED TO SEXUALITY? Y N

If YES, please indicate:

- | | | |
|--|---------|---------|
| 1. The wish to mess and soil. | 0 1 2 3 | 0 1 2 3 |
| 2. Wishing to acquire the unique affection of the parent of the opposite sex by displacing parent of the same sex (oedipus). | 0 1 2 3 | 0 1 2 3 |
| 3. Wish to be the sexual partner of the parent of the same sex. | 0 1 2 3 | 0 1 2 3 |
| 4. Other childhood sexual urges. Please specify: | 0 1 2 3 | 0 1 2 3 |
| <hr/> | | |
| 5. Fear of retaliatory punishment (e.g. castration). | 0 1 2 3 | 0 1 2 3 |
| 6. Masochistic pleasure at being hurt. | 0 1 2 3 | 0 1 2 3 |
| 7. Wish to be the sexual partner of the parent of the opposite sex (negative oedipus). | 0 1 2 3 | 0 1 2 3 |
| 8. Sexual over-stimulation and confusion. | 0 1 2 3 | 0 1 2 3 |
| 9. Fear of being excluded from a "sexual" couple. | 0 1 2 3 | 0 1 2 3 |
| 10. Fear of being ganged up on by a "sexual" couple. | 0 1 2 3 | 0 1 2 3 |
| 11. Fear of losing primary object through couple's sexual interest. | 0 1 2 3 | 0 1 2 3 |
| 12. Fear of merger (loss of self) in sexual union. | 0 1 2 3 | 0 1 2 3 |

Presence (PRES): 0 = absent
 1 = possibly present
 2 = present
 3 = prominent

Interpretation (INT): 0 = not taken up
 1 = taken into account in other interpretation
 2 = taken up directly
 3 = a central interpretation of the week

13. Omnipotent control over object through sexuality. 0 1 2 3 0 1 2 3

D. ARE THESE RELATED TO SELF AND SELF ESTEEM? Y N

If YES, please indicate

- 1. Unrealistically high internal standards. 0 1 2 3 0 1 2 3
- 2. Inconsistencies in the patient's internal standards. 0 1 2 3 0 1 2 3
- 3. About bolstering fragile sense of self (identity). 0 1 2 3 0 1 2 3
- 4. Dependence on object for self-esteem. 0 1 2 3 0 1 2 3
- 5. Fear of being overwhelmed (annihilated). 0 1 2 3 0 1 2 3
- 6. Identification with idealised part-object representation. 0 1 2 3 0 1 2 3
- 7. Internalisation of object's criticism. 0 1 2 3 0 1 2 3
- 8. About identification with denigrated part-object. 0 1 2 3 0 1 2 3
- 9. Object is obliterated (narcissistic stance). 0 1 2 3 0 1 2 3
- 10. Narcissistic grandiosity. 0 1 2 3 0 1 2 3

E. UNCONSCIOUS CONTENT RELATES TO BODY. Y N

If YES, please indicate:

- 1. Being in the womb. 0 1 2 3 0 1 2 3
- 2. Body as a phallus. 0 1 2 3 0 1 2 3
- 3. Identification with faeces. 0 1 2 3 0 1 2 3
- 4. Phallic display of body. 0 1 2 3 0 1 2 3
- 5. Other (please specify): 0 1 2 3 0 1 2 3

F. PREDOMINANT REACTION OF PATIENT TO INTERPRETATION

(Both 1 and 2 may apply)

1. Patient clearly feels helped. Y N

If YES, then indicate likely reason:

- a. Lowered anxiety. 0 1 2 3 0 1 2 3
- b. Feels understood. 0 1 2 3 0 1 2 3
- c. New insight. 0 1 2 3 0 1 2 3
- d. Reassurance. 0 1 2 3 0 1 2 3

2. Patient reports not feeling helped. Y N

If YES, then indicate likely reason.

- a. Patient does not consciously feel understood. 0 1 2 3 0 1 2 3
- b. Patient rejects UCS aspect of interpretations. 0 1 2 3 0 1 2 3
- c. Experiences interpretation as attack. 0 1 2 3 0 1 2 3

3. Analyst feels patient was helped. Y N

If YES, likely reason:

- a. Patient feels contained. 0 1 2 3 0 1 2 3
- b. Patient gains insight. 0 1 2 3 0 1 2 3
- c. Patient confronted with unpleasant internal reality. 0 1 2 3 0 1 2 3
- d. Patient confronted with unpleasant external reality. 0 1 2 3 0 1 2 3

Presence (PRES): 0 = absent
1 = possibly present
2 = present
3 = prominent

Interpretation (INT): 0 = not taken up
1 = taken into account in other interpretation
2 = taken up directly
3 = a central interpretation of the week

- | | | |
|---|---------|---------|
| e. Reassured of analyst's presence/involvement. | 0 1 2 3 | 0 1 2 3 |
| f. Other (please specify): | 0 1 2 3 | 0 1 2 3 |
-

4. Analyst feels patient reacts negatively to analyst

Y N

If YES, likely reason:

- | | | |
|--------------------------------------|---------|---------|
| a. Envious reaction. | 0 1 2 3 | 0 1 2 3 |
| b. Overwhelming guilt. | 0 1 2 3 | 0 1 2 3 |
| c. Overwhelming anxiety. | 0 1 2 3 | 0 1 2 3 |
| d. Misplaced, mistimed intervention. | 0 1 2 3 | 0 1 2 3 |
| e. Sado-masochistic relationship. | 0 1 2 3 | 0 1 2 3 |
| f. Rejection of dependency. | 0 1 2 3 | 0 1 2 3 |
| g. Other (please specify): | 0 1 2 3 | 0 1 2 3 |
-

G. ANALYST'S STYLE OF INTERVENTION

Indicate which aspect of your work had particular importance during this period.

Please use your own judgement on a scale 0 - 5, using the definitions in footnote as a guide.

- | | |
|---|-------------|
| 1. Acknowledging the patient's struggle to control his impulses. | 0 1 2 3 4 5 |
| 2. Provides support. | 0 1 2 3 4 5 |
| 3. Strengthening defences (e.g. support or agreement with splitting or denial). | 0 1 2 3 4 5 |
| 4. Asking for more information or elaboration. | 0 1 2 3 4 5 |
| 5. Comments on consistent quality in the patient's thoughts. | 0 1 2 3 4 5 |
| 6. Challenging the patient's view of some aspect of reality. | 0 1 2 3 4 5 |
| 7. Remarking on habitual aspects of the patient's behaviour (e.g. the patient's habit of withdrawing from/avoiding threatening situations). | 0 1 2 3 4 5 |
| 8. Relation to actual people in the patient's life. | 0 1 2 3 4 5 |
| 9. Connections between the therapeutic relationship and past relationships. | 0 1 2 3 4 5 |
| 10. Connections between the therapeutic relationship and present relationships. | 0 1 2 3 4 5 |
| 11. Comments about displacing feelings from the analyst to outside figure (lateral transference). | 0 1 2 3 4 5 |
| 12. Comments about patient's experiences of little child or baby within. | 0 1 2 3 4 5 |
| 13. Links between behaviour outside the session to current treatment material. | 0 1 2 3 4 5 |
| 14. Comments about possible meaning of others' behaviour. | 0 1 2 3 4 5 |
| 15. Comments on use of defensive manoeuvres, e.g. undoing, denial. | 0 1 2 3 4 5 |
| 16. The analyst comments on the patient's resistance (e.g. fear of change). | 0 1 2 3 4 5 |
| 17. Comments on patient's relationship to analytic process (e.g. attacking) | 0 1 2 3 4 5 |
| 18. Analyst writes letter or makes phone call to patient. | 0 1 2 3 4 5 |

Presence (PRES): 0 = absent
 1 = possibly present
 2 = present
 3 = prominent

Interpretation (INT): 0 = not taken up
 1 = taken into account in other interpretation
 2 = taken up directly
 3 = a central interpretation of the week

H. PREDOMINANT ASPECTS OF ANALYST'S FEELINGS

- | | |
|---|-----|
| 1. Patient's material triggers no feeling. | Y N |
| 2. Patient's material triggers no feeling where it would be expected. | Y N |
| 3. Patient's material triggers feelings in acceptable range i.e. useable counter transference feelings. | Y N |
| 4. Patient's material triggers analyst's emotional conflicts which intrude into the relationship. | Y N |

If YES to (3) or (4), then specify:

- | | |
|---|-------------|
| 5. Feels protective about patient. | 0 1 2 3 4 5 |
| 6. Feels loving and warm toward patient. | 0 1 2 3 4 5 |
| 7. Feels sensitive to the patient's feelings, attuned to the patient; empathic. | 0 1 2 3 4 5 |
| 8. Feels confident or self-assured. | 0 1 2 3 4 5 |
| 9. Feels angry and hostile. | 0 1 2 3 4 5 |
| 10. Feels belittled, denigrated. | 0 1 2 3 4 5 |
| 11. Feels persecuted and attacked. | 0 1 2 3 4 5 |
| 12. Feels frightened by the patient. | 0 1 2 3 4 5 |
| 13. Feels put upon, taken advantage of. | 0 1 2 3 4 5 |
| 14. Feels diffident and uncertain. | 0 1 2 3 4 5 |
| 15. Feels distant, aloof, uninvolved. | 0 1 2 3 4 5 |
| 16. Feels inadequate (not good enough). | 0 1 2 3 4 5 |
| 17. Feels tactless or insensitive. | 0 1 2 3 4 5 |
| 18. Feels bored and disinterested. | 0 1 2 3 4 5 |
| 19. Feels frustrated. | 0 1 2 3 4 5 |
| 20. Feels confused. | 0 1 2 3 4 5 |
| 21. Feels anxious about patient's self-destructive potential. | 0 1 2 3 4 5 |
| 22. Feels disgusted by patient. | 0 1 2 3 4 5 |

I. MAIN IMPACT OF PATIENT

- | | |
|---|-----|
| 1. Often the analyst thinks the main impact of the patient is not that primarily conveyed by the manifest content. Was this the case? | Y N |
|---|-----|

If YES, please indicate:

- | | |
|--|-------------|
| 2. Patient attempts to get rid of mental contents. | 0 1 2 3 4 5 |
| 3. Patient is in a bullying relationship with analyst. | 0 1 2 3 4 5 |
| 4. Patient attempts to shock/disgust analyst. | 0 1 2 3 4 5 |
| 5. Patient attempts to terrorise analyst. | 0 1 2 3 4 5 |
| 6. Patient attempts to set analyst up as a critic or external primitive superego. | 0 1 2 3 4 5 |
| 7. Patient is gratified by analytic interventions that are experienced as attacks or criticisms. | 0 1 2 3 4 5 |
| 8. Analyst placed in position of having exaggerated good qualities and value to patient. | 0 1 2 3 4 5 |
| 9. Analyst placed in position of having little of value to offer patient. | 0 1 2 3 4 5 |
| 10. Analyst is made to feel rejected while constantly reassured about importance. | 0 1 2 3 4 5 |

Please use your own judgement, using these definitions as a guide:

- | | |
|-----------------------------|-------------------------|
| 0 = Not applicable | 3 = Important |
| 1 = Not important or absent | 4 = Very important |
| 2 = Unimportant but present | 5 = Extremely important |

11. Other (please specify):

0 1 2 3 4 5

V. ANALYST'S JUDGEMENT OF THE QUALITY OF THE WEEK (TO BE RATED RELATIVE TO THE 'AVERAGE' PATIENT IN AN 'AVERAGE' WEEK).

VP=Very Poor. P=Poor. A=Average. G=Good. VG=Very Good.

- | | | | | | |
|---|-----------|----------|----------|----------|-----------|
| 1. Patient's general stance (therapeutic alliance). | VP | P | A | G | VG |
| 2. Quality of patient's manifest material. | VP | P | A | G | VG |
| 3. Completeness and quality of analyst's understanding of patient's material. | VP | P | A | G | VG |
| 4. Quality of analyst's interpretive interventions. | VP | P | A | G | VG |
| 5. The quality of the patient's response to interpretation. | VP | P | A | G | VG |
| 6. Impression of overall quality of week. | VP | P | A | G | VG |

Please use your own judgement, using these definitions as a guide:

0 = Not applicable

1 = Not important or absent

2 = Unimportant but present

3 = Important

4 = Very important

5 = Extremely important

APPENDIX 4.2A. RESULTS OF YAWRS SUBSECTION FACTOR ANALYSES.

Note: Each numbered section of the table below lists the YAWRS items included in a single factor analysis, along with the loadings of these items on 1 to 5 derived factors, the labels given to these factors, the variance accounted for by each factor, the total variance accounted for by the entire model, and the number of YAWRS forms used in the calculation (limited by the number of forms in which the items in the appropriate subsection were filled out). Factors were calculated by summing the items corresponding to loadings in bold, with weights of only +1 or -1. Cronbach's alpha was calculated for each calculated factor consisting of more than one item.

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
<i>0. Sample</i>		Title	Title			
	Variance	Variance	Variance			
	n	α	α			
Variable 1		loading	loading			
Variable 2		loading	loading			
<i>1. Resistance – general (GA1)</i>	0.38	0.38				
	958	0.66				
GA01		-0.84				
GA02		0.53				
GA03		0.68				
GA04		0.22				
<i>2. Resistance-examples (GA2)</i>		Active	Passive			
	0.33	0.21	0.12			
	529	0.74	0.51			
GA05p		-0.10	0.72			
GA06p		0.49	0.31			
GA07p		0.42	0.17			
GA08p		0.41	0.03			
GA09p		0.73	-0.04			
GA10p		0.10	0.22			
GA11p		-0.03	0.63			
GA12p		0.68	0.19			
GA13p		0.63	-0.07			
GA14p		-0.23	-0.34			

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
3. General attitude – negative (GB)	0.25 958	0.25 0.63				
GB01p		-0.50				
GB02p		0.58				
GB03p		0.47				
GB04p		0.26				
GB05p		0.58				
GB06p		0.51				
4. Time keeping (GC)		Difficulty leaving	Other time problems			
	0.3 518	0.18 0.52	0.12 0.31			
GC02p		0.06	0.44			
GC03p		-0.65	-0.31			
GC04p		0.68	-0.19			
GC05p		0.05	0.44			
GC06p		-0.05	0.26			
5. Bad behaviour in analysis (GE)	0.11 478	0.11 0.23				
GE02p		0.72				
GE03p		-0.14				
GE04p		0.04				
GE05p		0.03				
GE06p		0.61				
GE07p		0.35				
GE08p		0.00				
GE09p		-0.02				
GE10p		0.25				
GE11		0.18				
6. Quality of analytic material (GF)	0.32 814	0.32 0.74				
GF02p		0.39				
GF03p		0.68				
GF04p		0.57				
GF05p		0.26				

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
GF06p		0.54				
GF07p		0.68				
GF08p		0.71				
7. Acting out: aggression and sexuality (GG)	0.15 151	0.15 -1.32				
GG02p		-0.87				
GG03p		0.28				
GG04p		0.17				
GG05p		0.21				
GG06p		0.53				
GG07p		0.09				
GG08p		-0.42				
GG09p		0.02				
GG10p		0.07				
8. Maturity of mental functioning (GH2)			Confusion/ Regression concrete	Primitive boundaries		
	0.42 809	0.17 0.64	0.16 0.64	0.10 0.57		
GH05p		0.33	0.50	0.07		
GH06p		0.22	0.13	0.66		
GH07p		0.32	0.36	0.41		
GH08p		0.19	0.68	0.09		
GH09p		-0.05	0.57	0.36		
GH10p		0.53	0.08	0.17		
GH11p		0.58	0.18	0.22		
GH12p		0.65	0.13	0.06		
9. Body pride/disgust (MA2)			Pride	Disgust		
	0.30 714	0.22 0.69	0.08			
MA04		0.74	-0.11			
MA05		0.06	-0.16			
MA06		0.12	0.54			
MA07		0.74	0.24			
MA08		0.09	-0.02			

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
<i>10. Body topics (MA3)</i>		Psychotic	Damaged	Narcis- sistic		
	0.36	0.16	0.11	0.09		
	714	0.57	0.62	0.47		
MA09p		0.25	0.38	0.18		
MA10p		0.22	0.64	0.03		
MA11p		0.12	0.00	0.55		
MA12p		0.36	0.03	0.24		
MA13p		0.29	0.29	0.52		
MA14p		0.70	0.00	0.12		
MA15p		0.48	0.54	0.26		
MA16p		0.58	0.30	0.26		
MA17p		-0.04	0.12	0.00		
<i>11. Self-esteem (MB)</i>		Positive	Negative			
	0.42	0.24	0.18			
	856	0.71	0.62			
MB02p		0.12	0.68			
MB03p		0.56	0.26			
MB04p		-0.09	0.54			
MB05p		0.66	-0.06			
MB06p		0.48	0.24			
MB07p		0.82	-0.14			
MB08p		0.09	0.47			
<i>12. Affective tone of memories (MC04)</i>		Anxious	Angry			
	0.24	0.13	0.11			
	340	0.57	0.68			
MC04Ap		-0.17	-0.12			
MC04Bp		0.19	0.16			
MC04Cp		0.38	0.12			
MC04Dp		0.70	0.12			
MC04Ep		0.23	0.67			
MC04Fp		0.29	0.26			
MC04Gp		0.18	0.08			
MC04Hp		0.10	0.73			

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
MC04Ip		0.60	0.02			
MC04Jp		0.48	0.18			
MC04Kp		-0.06	-0.25			
<i>13. Relationship themes w/family (MD04)</i>		Identifica- tion/indep -endence	Punish- ment	Narcissism	Oedipal	
	0.35	0.10	0.09	0.08	0.08	
	570	0.66	0.68	0.67	0.65	
MD04Ap		0.10	0.04	0.62	0.07	
MD04Bp		0.15	0.57	0.24	0.04	
MD04Cp		0.27	0.19	0.30	0.00	
MD04Dp		0.44	0.15	0.19	0.21	
MD04Ep		0.39	0.08	0.25	0.19	
MD04Fp		0.40	0.10	0.49	0.07	
MD04Gp		0.19	0.07	0.49	0.31	
MD04Hp		0.30	0.08	0.04	0.10	
MD04Ip		0.24	0.16	0.14	0.63	
MD04Jp		0.14	0.10	0.46	0.34	
MD04Kp		0.60	0.09	0.12	0.27	
MD04Lp		0.59	0.15	0.18	0.08	
MD04Mp		0.13	0.05	0.11	0.57	
MD04Np		0.27	0.39	0.14	0.43	
MD04Op		0.18	0.67	0.02	0.05	
MD04Pp		0.14	0.63	0.08	0.11	
MD04Qp		0.08	-0.14	0.06	0.17	
<i>14. Relationship themes w/friends (ME04)</i>		Narcissism	Belonging to group	Threatened	Separate- ness	
	0.45	0.14	0.11	0.10	0.09	
	489	0.66	0.79	0.73	0.72	
ME04Ap		0.22	0.16	0.67	0.08	
ME04Bp		0.25	0.17	0.10	0.14	
ME04Cp		0.20	0.13	0.74	0.15	
ME04Dp		0.60	0.05	0.31	-0.04	
ME04Ep		0.33	0.15	0.31	0.40	
ME04Fp		0.38	0.10	0.13	0.71	

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
ME04Gp		0.52	0.22	0.16	0.54	
ME04Hp		0.68	0.21	0.09	0.20	
ME04Ip		0.57	0.11	0.18	0.22	
ME04Jp		0.04	0.21	0.08	0.32	
ME04Kp		0.14	0.23	0.39	0.31	
ME04Lp		0.14	0.71	0.15	0.09	
ME04Mp		0.38	0.62	0.18	0.04	
ME04Np		0.18	0.71	0.21	0.23	
ME04Op		-0.02	-0.07	-0.15	0.16	
15. Relationship themes w/sexual partners (MF04)			Fear of being unloved	Fear of attack	Wish for stability	
	0.45	0.16	0.14	0.08	0.07	
	358	0.75	0.79	0.72	0.63	
MF04Ap		0.16	0.09	0.75	-0.02	
MF04Bp		0.37	0.40	0.06	0.18	
MF04Cp		0.79	0.05	0.12	0.05	
MF04Dp		0.33	0.15	0.65	-0.03	
MF04Ep		0.06	0.78	0.16	-0.05	
MF04Fp		0.78	0.04	0.17	0.03	
MF04Gp		0.53	0.20	0.14	0.05	
MF04Hp		0.07	0.79	0.13	0.06	
MF04Ip		0.28	0.69	0.15	0.04	
MF04Jp		0.30	0.16	-0.03	0.03	
MF04Kp		0.43	0.30	0.09	-0.07	
MF04Lp		0.09	0.16	0.06	0.68	
MF04Mp		0.02	-0.09	0.04	0.70	
MF04Np		0.49	-0.06	0.28	0.08	
MF04Op		-0.01	-0.09	-0.22	-0.09	
16. Money and work (MG)	0.23	0.23				
	540	0.57				
MG02p		0.75				
MG03p		-0.02				
MG04p		0.17				
MG05p		0.02				

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
MG06p		-0.18				
MG07p		0.84				
MG08p		0.54				
<i>17. Affective reaction to event (MH05)</i>		Full of affect	Depressed/ suicidal	Not happy		
	0.42	0.18	0.15	0.09		
	475	0.54	0.50			
MH05Ap		0.03	-0.26	-0.65		
MH05Bp		0.48	0.26	0.00		
MH05Cp		0.36	0.61	0.20		
MH05Dp		0.18	0.61	0.02		
MH05Ep		0.48	0.15	0.23		
MH05Fp		-0.67	0.05	0.24		
<i>18. Gender, age, race (MI02)</i>		Patient culture/race	Analyst culture/race	Analyst gender	Analyst age	
	0.24	0.13	0.11			
	238	0.28	0.35	-0.06		
MI02Ap		-0.50	-0.03	1		
MI02Bp		-0.17	-0.01	1		
MI02Cp		-0.16	0.27		1	
MI02Dp		0.09	0.52			
MI02Ep		0.62	0.21			
MI02Fp		0.11	0.61			
MI02Gp		0.46	-0.12			
<i>19. Adult identity (MJ02)</i>	0.17	0.17				
	285	0.30				
MJ02Ap		0.18				
MJ02Bp		-0.11				
MJ02Cp		0.14				
MJ02Dp		0.68				
MJ02Ep		0.57				

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
<i>20. Sexual material – content (MK05)</i>			Sexual fantasy/ homosex	Sexual life		
		0.32	0.20	0.08	0.07	
		544	0.83	0.53	-0.59	
	MK05Ap		-0.17	0.49	0.07	
	MK05Bp		0.16	0.26	0.62	
	MK05Cp		0.10	0.23	0.05	
	MK05Dp		-0.02	-0.05	0.68	
	MK05Ep		0.43	0.54	0.17	
	MK05Fp		0.08	0.57	0.11	
	MK05Gp		0.07	-0.07	0.09	
	MK05Hp		0.01	0.36	-0.15	
	MK05Ip		-0.01	-0.19	0.11	
	MK05J1p		0.85	-0.10	0.03	
	MK05J2p		0.86	-0.08	-0.04	
	MK05J3p		0.76	0.06	0.13	
	MK05J4p		0.59	0.22	0.01	
MK05J5p		0.55	0.09	0.10		
MK05Kp		0.11	0.09	0.02		
<i>21. Sexual material – motivation (MK06)</i>			Inhibition /anxiety	Shame/am- bivalence		
		0.34	0.21	0.13		
		544	0.64	0.53		
	MK06Ap		0.73	0.15		
	MK06Bp		0.49	0.14		
	MK06Cp		0.56	0.19		
	MK06Dp		0.43	0.55		
MK06Ep		0.07	0.60			
MK06Fp		0.12	0.12			
<i>22. Affect discussing tx parameters (ML03)</i>			Sad/angry	Guilty/ anxious		
		0.34	0.20	0.14		
		712	0.69	0.71		
	ML03Ap		-0.06	0.00		
ML03Bp		0.39	0.23			

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
ML03Cp		0.84	0.09			
ML03Dp		0.68	0.12			
ML03Ep		0.24	0.63			
ML03Fp		0.42	0.57			
ML03Gp		0.13	0.67			
ML03Hp		-0.45	0.10			
ML03Ip		-0.04	0.08			
<i>23. General transference(PA)</i>	0.15 958	0.15 0.44				
PA01Ap		0.05				
PA01Bp		0.36				
PA01Cp		0.36				
PA02		-0.10				
PA03		0.20				
PA04		0.64				
PA05		0.56				
PA06		0.41				
<i>24. Positive transference (PA02)</i>			Love/ erotic	Idealiza- tion/iden- tification	Depen- dence	
	0.45 633	0.20 0.76		0.13 0.62	0.11	
PA02Ap		0.39		0.27	0.18	
PA02Bp		0.43		0.42	0.24	
PA02Cp		0.48		0.20	0.47	
PA02Dp		0.69		0.09	0.21	
PA02Ep		0.13		0.50	0.42	
PA02Fp		0.06		0.19	0.56	
PA02Gp		0.34		0.61	0.25	
PA02Hp		0.64		0.33	0.04	
<i>25. Transference w/anxiety (PA03)</i>			Paranoid	Fear of rejection	Projected aggression	
	0.5 794	0.19 0.72		0.16 0.74	0.14 0.72	
PA03Ap		0.48		0.16	0.30	
PA03Bp		0.24		0.62	0.27	

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
PA03Cp		0.38	0.17	0.67		
PA03Dp		0.77	0.12	0.15		
PA03Ep		0.59	0.28	0.21		
PA03Fp		0.16	0.75	0.16		
PA03Gp		0.41	0.50	0.16		
PA03Hp		0.16	0.27	0.69		
PA03Ip		0.41	0.22	0.20		
<hr/>						
<i>26. Transference w/competition and aggression (PA04)</i>		Analyst is helpless	Rivalry, victory			
	0.22	0.11	0.11			
	457	0.54	0.47			
PA04Ap		0.18	0.43			
PA04Bp		0.08	0.66			
PA04Cp		0.05	0.21			
PA04Dp		0.09	0.16			
PA04Ep		0.10	-0.16			
PA04Fp		0.69	0.07			
PA04Gp		0.51	0.15			
<hr/>						
<i>27. Transference w/resentment (PA05)</i>		Derogator y to analyst	Aban- doned by analyst	Loss of contact	Depen- dence	
	0.35	0.19	0.07	0.07	0.03	
	648	0.79	0.36	0.38		
PA05Ap		0.09	0.24	0.01	0.31	
PA05Bp		0.12	0.14	0.49	0.08	
PA05Cp		-0.02	0.47	0.08	0.11	
PA05Dp		0.66	0.19	0.17	-0.15	
PA05Ep		0.80	-0.12	0.24	0.07	
PA05Fp		0.75	-0.19	0.22	0.35	
PA05Gp		0.04	0.14	0.10	0.00	
PA05Hp		0.20	-0.06	0.45	-0.04	
PA05Ip		-0.01	0.48	0.00	0.01	

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
28. Primitive transference (PA06)	0.11 487	0.11 0.25				
PA06Ap		0.27				
PA06Bp		0.16				
PA06Cp		0.35				
PA06Dp		0.50				
PA06Ep		0.43				
PA06Fp		-0.18				
29. Sadness (PC02)	0.23 472	0.23 0.53				
PC02Ap		0.10				
PC02Bp		0.56				
PC02Cp		0.63				
PC02Dp		0.63				
PC02Ep		-0.09				
30. Primitive emotional stance (PC03)			Fear/ shame	Manic grandiosity	Existential anxiety	
	0.42 800	0.20 0.77	0.12 0.71	0.10 0.57		
PC03Ap		0.21	-0.08	0.40		
PC03Bp		0.59	-0.01	0.11		
PC03Cp		0.41	0.18	0.18		
PC03Dp		0.20	0.66	0.04		
PC03Ep		0.03	0.81	-0.06		
PC03Fp		-0.17	0.38	0.02		
PC03Gp		0.08	-0.01	0.76		
PC03Hp		0.34	0.23	0.46		
PC03Ip		0.72	0.01	0.15		
PC03Jp		0.72	-0.03	0.11		
PC03Kp		0.65	-0.08	0.28		
31. Anger (PC04)	0.17 660	0.17 0.40				
PC04Ap		0.30				
PC04Bp		0.41				

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
PC04Cp		-0.60				
PC04Dp		0.27				
<i>32. Predominant defences (PD)</i>		Projection, reaction formation, denial	Isolation	Splitting	Regression, projective identifica- tion	Exter- nali- sation
	0.40	0.13	0.09	0.07	0.07	0.05
	958	0.70	0.62	0.54	0.53	
PD01p		-0.18	0.66	0.10	0.05	0.16
PD02p		0.07	0.13	0.02	0.07	0.69
PD03p		0.44	0.01	0.13	0.24	0.35
PD04p		-0.11	0.01	-0.09	0.59	0.09
PD05p		0.59	-0.01	-0.01	0.02	0.01
PD06p		0.57	0.30	0.01	-0.15	0.09
PD07p		0.10	0.10	0.10	0.09	-0.09
PD08p		0.03	0.39	0.22	0.47	0.07
PD09p		0.11	0.15	0.58	0.14	0.15
PD10p		0.30	0.17	0.62	-0.10	0.17
PD11p		0.53	0.32	0.24	0.22	-0.10
PD12p		0.25	0.68	0.07	0.12	0.03
PD13p		0.65	-0.08	0.19	0.01	0.06
PD14p		0.33	0.15	0.14	0.49	-0.02
PD15p		-0.02	-0.01	0.37	0.01	-0.09
<i>33. UCS general (UA)</i>		Narcissis- tic infant objects	Valued object	Greed/ envy		
	0.39	0.16	0.15	0.09		
	887	0.55	0.68			
UA03p		0.33	0.60	0.13		
UA04p		0.17	0.14	0.59		
UA05p		0.43	0.46	0.16		
UA06p		0.54	0.21	0.18		
UA07p		0.18	0.53	0.43		
UA08p		0.30	0.25	0.11		
UA09p		0.61	0.21	0.12		

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
34. UCS reactions to aggression (UB)		Destruc- tive	Anxious	Maso- chistic	Depression	
	0.43	0.14	0.13	0.11	0.06	
	766	0.71	0.72	0.64		
UB01p		-0.11	0.23	0.29	0.19	
UB02p		0.52	0.05	0.47	0.10	
UB03p		0.61	0.17	0.30	0.19	
UB04p		0.49	0.03	0.04	-0.02	
UB05p		0.13	0.63	0.25	-0.02	
UB06p		0.11	0.62	0.07	0.28	
UB07p		0.09	0.72	0.25	-0.01	
UB08p		0.61	0.16	0.17	-0.12	
UB09p		0.56	0.03	-0.15	0.13	
UB10p		0.06	0.06	0.14	0.67	
UB11p		0.15	0.19	0.57	0.20	
UB12p		0.10	0.29	0.63	-0.01	
35. UCS related to sexuality (UC)		Oedipal	Identity	Maso- chism		
	0.38	0.19	0.10	0.08		
	547	0.80	0.61	0.50		
UC01p		0.37	0.05	0.47		
UC02p		0.55	0.23	0.46		
UC03p		0.36	0.05	0.02		
UC04p		0.02	0.01	-0.13		
UC05p		0.39	0.21	0.33		
UC06p		0.20	0.05	0.40		
UC07p		0.45	0.19	0.38		
UC08p		0.00	0.57	-0.22		
UC09p		0.81	-0.01	0.04		
UC10p		0.52	0.25	0.12		
UC11p		0.73	0.15	0.24		
UC12p		0.12	0.55	0.30		
UC13p		0.29	0.65	0.13		

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
36. UCS related to self and self-esteem (UD)		Low self-esteem	Narcissistic	Obliterated	Self-criticism	
	0.53	0.18	0.13	0.13	0.10	
	815	0.77	0.70	0.67	0.61	
UD01p		0.27	-0.02	0.21	0.66	
UD02p		0.40	0.14	0.24	0.42	
UD03p		0.14	-0.02	0.74	0.09	
UD04p		0.39	-0.09	0.47	0.11	
UD05p		0.14	0.17	0.59	0.30	
UD06p		0.67	0.35	0.07	0.31	
UD07p		0.57	-0.03	0.16	0.35	
UD08p		0.74	0.08	0.21	0.11	
UD09p		-0.07	0.79	0.10	-0.02	
UD10p		0.17	0.70	-0.06	0.06	
37. UCS related to body (UE)	0.28	0.28				
	466	0.35				
UE01p		0.24				
UE02p		0.86				
UE03p		0.18				
UE04p		0.72				
UE05p		-0.23				
38. Patient feels helped by interpretation (UF01)	0.61	0.61				
	516	0.86				
UF01Ap		0.82				
UF01Bp		0.94				
UF01Cp		0.66				
UF01Dp		0.69				
39. Patient feels not helped by interpretation (UF02)	0.25	0.25				
	182	0.40				
UF02Ap		0.29				
UF02Bp×2		0.70				
UF02Cp×3		0.41				

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
<i>40. Analyst feels patient helped by interpretation (UF03)</i>			Contained	Confronted		
	0.35	0.19	0.16			
	700	0.66	0.59			
	UF03Ap	0.73	0.10			
	UF03Bp	0.29	0.38			
	UF03Cp	0.22	0.74			
	UF03Dp	0.13	0.49			
	UF03Ep	0.67	0.06			
UF03Fp	-0.01	0.02				
<i>41. Analyst feels patient reacted negatively to interpretation (UF04)</i>	0.14	0.14				
	514	0.31				
	UF04Ap	-0.13				
	UF04Bp	0.38				
	UF04Cp	0.81				
	UF04Dp	0.35				
	UF04Ep	0.07				
	UF04Fp	0.13				
UF04Gp	-0.09					
<i>42. Analyst's style of intervention (UG)</i>		Interprets object relations	Interprets thinking, perception, behaviour	Supportive	Interprets defences	
	0.38	0.11	0.10	0.09	0.08	
	958	0.75	0.73	0.71	0.61	
	UG01	0.06	0.29	0.36	0.22	
	UG02	0.24	0.20	0.64	0.04	
	UG03	0.03	0.17	0.55	0.08	
	UG04	-0.05	0.20	0.18	0.33	
	UG05	0.10	0.61	0.16	0.15	
UG06	0.06	0.69	0.05	0.23		
UG07	0.07	0.45	0.14	0.45		

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
UG08		0.50	0.00	0.40	-0.05	
UG09		0.54	0.03	-0.06	0.05	
UG10		0.68	0.34	0.27	0.04	
UG11		0.48	0.43	0.16	0.04	
UG12		0.48	-0.18	0.03	0.14	
UG13		0.56	0.26	0.21	0.04	
UG14		0.01	0.16	0.58	0.06	
UG15		0.18	0.10	0.06	0.62	
UG16		0.14	0.06	-0.02	0.71	
UG17		0.27	0.22	0.06	0.34	
UG18		0.05	-0.09	0.35	0.05	
43. Analyst's feelings (UH)			Inadequate /confused	Loving/ empathic	Bored/ cutoff	
	0.48	0.16	0.14	0.10	0.08	
	910	0.81	0.76	0.78	0.67	
UH05		-0.01	0.37	0.58	-0.15	
UH06		-0.16	0.06	0.74	-0.01	
UH07		-0.01	-0.04	0.73	-0.14	
UH08		-0.07	-0.13	0.72	0.13	
UH09		0.69	0.22	-0.13	0.19	
UH10		0.77	0.18	-0.16	0.11	
UH11		0.71	0.23	-0.02	-0.02	
UH12		0.33	0.31	0.22	0.08	
UH13		0.62	0.11	-0.30	0.20	
UH14		0.05	0.63	-0.07	0.28	
UH15		0.13	0.20	0.00	0.71	
UH16		0.24	0.67	0.01	0.03	
UH17		0.26	0.53	0.14	0.23	
UH18		0.13	0.07	-0.14	0.63	
UH19		0.46	0.30	-0.33	0.14	
UH20		0.25	0.60	-0.06	0.12	
UH21		0.23	0.54	0.00	-0.01	
UH22		0.39	0.29	0.09	0.31	

	<i>Total</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>5</i>
44. Impact of patient (UI)			Narcis- sistic	Analyst rejected		
	0.33	0.17	0.10	0.06		
	567	0.70	0.61			
UI02		0.30	0.05	0.21		
UI03		0.68	0.26	0.17		
UI04		0.61	0.09	0.02		
UI05		0.66	0.27	-0.12		
UI06		0.19	0.45	0.04		
UI07		0.36	0.58	-0.09		
UI08		0.13	0.46	-0.23		
UI09		0.19	0.10	0.53		
UI10		0.03	0.54	0.06		
UI11		0.07	0.20	-0.33		
45. Quality of week (J)	0.61	0.61				
	958	0.90				
J1		0.79				
J2		0.73				
J3		0.68				
J4		0.68				
J5		0.84				
J6		0.94				

APPENDIX 4.2B. RESULTS OF SUBSECTION SUMMARY SCALE FACTOR ANALYSIS.

Total variance accounted for = 0.20 (n=958)

	<i>Resistance</i>	<i>Ucs themes & Regression</i>
Variance =	0.12	0.08
Alpha =	0.94	0.67
GAop	0.55	-0.37
GBop	0.52	-0.40
GCop	0.05	-0.21
GDop	0.01	-0.14
GEop	0.26	-0.08
GFop	-0.58	0.20
GGop	0.24	0.23
GHop	0.47	-0.12
MAop	-0.26	0.15
MBop	0.04	0.37
MCop	-0.24	0.04
MDop	-0.34	-0.10
MEop	-0.33	0.12
MFop	-0.07	0.19
MGop	-0.08	-0.04
MHop	-0.05	0.09
MIop	-0.08	0.35
MJop	-0.47	0.11
MKop	-0.09	0.37
MLop	-0.21	0.09
PA1op	-0.01	0.08
PA2op	-0.52	0.48
PA3op	0.04	0.56
PA4op	0.52	0.11
PA5op	0.51	0.10
PA6op	0.52	0.03
PBop	-0.08	0.08
PC1op	-0.42	-0.15
PC2op	-0.49	0.23

	<i>Resistance</i>	<i>Ucs themes & Regression</i>
PC3op	0.20	0.50
PC4op	0.22	0.33
PDf1p	0.05	-0.18
PDf2p	0.19	-0.38
PDf3p	0.16	0.12
PDf4p	0.35	-0.01
PDf5p	0.24	-0.39
UAop	-0.27	0.35
UBop	0.33	0.49
UCop	-0.01	0.64
UDop	-0.04	0.49
UEop	-0.03	0.49
UF1p	-0.66	0.29
UF2p	0.27	0.02
UF3p	-0.52	0.23
UF4p	0.54	0.03
UGf1o	-0.48	-0.14
UGf2o	-0.07	-0.46
UGf3o	-0.26	-0.20
UGf4o	0.03	-0.43
UHf1o	0.57	-0.18
UHf2o	0.29	-0.09
UHf3o	-0.67	0.01
UHf4o	0.35	-0.32
UIf1o	0.47	-0.03
UIf2o	0.39	-0.23
UIf3o	0.44	-0.32

APPENDIX 4.3A. FORMULAE FOR CALCULATION OF FACTORS FROM FACTOR ANALYSES (FAC).

Sample:

Subsection

1. Factor 1 description
Factor 1 variable name = ...; n=x,y
2. Factor 2 description
Factor 2 variable name = ...; n=x,y

where x = # of variables averaged using data from printed YAWRS forms D, E, and G and y = # of variables averaged using data from printed YAWRS forms C, C2, C3 (if different from x)

General Characteristics

GA1

1. Resistance - general description
GA1 = mean(1-GA01,GA02,GA03,GA04); n=4,3

GA2

2. Resistance: f1=active examples - pres
GA2f1p = mean(GA06p,GA07p,GA08p,GA09p,GA12p,GA13p); n=6
3. Resistance: f2=passive examples - pres
GA2f2p = mean(GA05p,GA10p,GA11p); n=3
4. Resistance: f1=active examples - int
GA2f1i = mean(GA06i,GA07i,GA08i,GA09i,GA12i,GA13i); n=6
5. Resistance: f2=passive examples - int
GA2f2i = mean(GA05i,GA10i,GA11i); n=3

GB

6. General negative attitude to analysis - pres
GBp = mean(3-GB01p,GB02p,GB03p,GB04p,GB05p,GB06p); n=6,4
7. General negative attitude to analysis - int
GBi = mean(GB01i,GB02i,GB03i,GB04i,GB05i,GB06i); n=6,4

GC

8. Problems with time keeping: f1=difficulty leaving - pres
GCf1p = mean(GC03p,GC04p); n=2
9. Problems with time keeping: f2=other time problems - pres
GCf2p = mean(GC02p,GC05p,GC06p); n=3
10. Problems with time keeping: f1=difficulty leaving - int
GCf1i = mean(GC03i,GC04i); n=2
11. Problems with time keeping: f2=other time problems - int
GCf2i = mean(GC02i,GC05i,GC06i); n=3

GD

12. Patterns and meaning in missed sessions - pres
GDp = mean(GD05p,GD06p); n=2,NA
13. Patterns and meaning in missed sessions - int
GDi = mean(GD05i,GD06i); n=2,NA

GE

14. Bad behaviour in session - pres
GEp = mean(GE02p,GE03p,GE04p,GE05p,GE06p,GE07p,GE08p,GE09p,GE10p,3*GE11); n=10,8

15. Bad behaviour in session - int
GEi = mean(GE02i,GE03i,GE04i,GE05i,GE06i,GE07i,GE08i,GE09i,GE10i); n=9,7

GF

16. Quality of analytic material - pres
GFp = mean(GF02p,GF03p,GF04p,GF05p,GF06p,GF07p,GF08p); n=7

17. Quality of analytic material - int
GFi = mean(GF02i,GF03i,GF04i,GF05i,GF06i,GF07i,GF08i); n=7

GG

18. Acting out: aggression and sexuality in analysis - pres
GGp = mean(GG02p,GG03p,GG04p,GG05p,GG06p,GG07p,GG08p,GG09p,GG10p); n=9,8

19. Acting out: aggression and sexuality in analysis - int
GGi = mean(GG02i,GG03i,GG04i,GG05i,GG06i,GG07i,GG08i,GG09i,GG10i); n=9,8

GH1

20. Immaturity of mental functioning (0,1,2)
GH1 = 1 if GH02 or GH03 is "yes", 2 if GH04 is yes

GH2

21. Maturity of mental functioning: f1=regression - pres
GH2f1p = mean(GH10p,GH11p,GH12p); n=3

22. Maturity of mental functioning: f2=confusion & concreteness - pres
GH2f2p = mean(GH05p,GH08p,GH09p); n=3

23. Maturity of mental functioning: f3=primitive boundary problems - pres
GH2f3p = mean(GH06p,GH07p); n=2

24. Maturity of mental functioning: f1=regression - int
GH2f1i = mean(GH10i,GH11i,GH12i); n=3

25. Maturity of mental functioning: f2=confusion & concreteness - int
GH2f2i = mean(GH05i,GH08i,GH09i); n=3

26. Maturity of mental functioning: f3=primitive boundary problems - int
GH2f3i = mean(GH06i,GH07i); n=2

Manifest Content

MA1

27. Role of body in material (0=absent, 1=occasional, 2=significant)
MA1 = 1 if MA02 is yes, 2 if MA03 is yes, 0 otherwise

MA2

28. Satisfaction or pride w/body
MA2f1 = mean(MA04,MA07); n=2,1

29. Dislike or disgust w/body
MA2f2 = MA06

MA3

30. Body factor: f1 = psychotic body - pres
MA3f1p = mean(MA12p,MA14p,MA16p); n=3,2

31. Body factor: f2 = damage to own body - pres
MA3f2p = mean(MA09p,MA10p,MA15p); n=3,1

32. Body factor: f3 = narcissistic body - pres
MA3f3p = mean(MA11p,MA13p); n=2

33. Body factor: f1 = psychotic body - int
MA3f1i = mean(MA12i,MA14i,MA16i); n=3,2

34. Body factor: f2 = damage to own body - int
MA3f2i = mean(MA09i,MA10i,MA15i); n=3,1

35. Body factor: f3 = narcissistic body - int
MA3f3i = mean(MA11i,MA13i); n=2

MB

36. Self-esteem: f1 = positive self-esteem - pres
MBf1p = mean(MB03p,MB05p,MB06p,MB07p); n=4,3
37. Self-esteem: f2 = negative self-esteem - pres
MBf2p = mean(MB02p,MB04p,MB08p); n=3,1
38. Self-esteem: f1 = positive self-esteem - int
MBf1i = mean(MB03i,MB05i,MB06i,MB07i); n=4,3
39. Self-esteem: f2 = negative self-esteem - int
MBf2i = mean(MB02i,MB04i,MB08i); n=3,1

MC02

40. Hist mat (0=none, 1=repert, 2=new perspec, 3=new material) - pres
MC02p = 1 if MC02Cp ≥ 2, 2 if MC02Bp ≥ 2, 3 if MC02Ap ≥ 2
41. Hist mat (0=none, 1=repert, 2=new perspec, 3=new material) - int
MC02i = 1 if MC02Ci ≥ 2, 2 if MC02Bi ≥ 2, 3 if MC02Ai ≥ 2

MC03 = time period of historical material

Years ago weighting:	young adulthood	= 3.5
	adolescence	= 8.5
	pre-adolescence	= 12
	junior school	= 15.5
	infant school	= 18.5
	early childhood	= 22
	infancy	= 24

42. Hist mat: total content weighted by 'years ago' (0 to 312) - pres
MC03ap = sum(MC03Ap*24,MC03Bp*22,MC03Cp*18.5,MC03Dp*15.5,
MC03Ep*12,MC03Fp*8.5,MC03Gp*3.5)
43. Hist mat: 'years ago' with maximum presence (0 to 24) - pres
MC03bp = years ago weighting corresponding to time of greatest pres score
44. Hist mat: oldest 'years ago' present (0 to 24) - pres
MC03cp = years ago weighting corresponding to oldest time with pres score > 0
45. Hist mat: total content weighted by 'years ago' (0 to 312) - int
MC03ai = sum(MC03Ai*24,MC03Bi*22,MC03Ci*18.5,MC03Di*15.5,
MC03Ei*12,MC03Fi*8.5,MC03Gi*3.5)
46. Hist mat: 'years ago' with maximum presence (0 to 24) - int
MC03ip = years ago weighting corresponding to time of greatest int score
47. Hist mat: oldest 'years ago' present (0 to 24) - int
MC03cp = years ago weighting corresponding to oldest time with int score > 0

MC04

48. Affective tone of memories: f1 = anxious - pres
MC04f1p = mean(MC04Cp,MC04Dp,MC04Ip,MC04Jp); n=4,2
49. Affective tone of memories: f2 = angry - pres
MC04f2p = mean(MC04Ep,MC04Hp); n=2,NA
50. Affective tone of memories: f1 = anxious - int
MC04f1i = mean(MC04Ci,MC04Di,MC04Ii,MC04Ji); n=4,2
51. Affective tone of memories: f2 = angry - int
MC04f2i = mean(MC04Ei,MC04Hi); n=2,NA

MD04

52. Relationship themes w/family: f1 = identification/independence - pres
MD04f1p = mean(MD04Dp,MD04Ep,MD04Hp,MD04Kp,MD04Lp); n=5,3
53. Relationship themes w/family: f2 = punishment - pres
MD04f2p = mean(MD04Bp,MD04Op,MD04Pp); n=3,1
54. Relationship themes w/family: f3 = narcissism - pres
MD04f3p = mean(MD04Ap,MD04Cp,MD04Fp,MD04Gp,MD04Jp); n=5,4
55. Relationship themes w/family: f4 = Oedipal - pres
MD04f4p = mean(MD04Ip,MD04Mp,MD04Np); n=3,1
56. Relationship themes w/family: f1 = identification/independence - int
MD04f1i = mean(MD04Di,MD04Ei,MD04Hi,MD04Ki,MD04Li); n=5,3

57. Relationship themes w/family: f2 = punishment - int
MD04f2i = mean(MD04Bi,MD04Oi,MD04Pi); n=3,1
58. Relationship themes w/family: f3 = narcissism - int
MD04f3i = mean(MD04Ai,MD04Ci,MD04Fi,MD04Gi,MD04Ji); n=5,4
59. Relationship themes w/family: f4 = Oedipal - int
MD04f4i = mean(MD04Ii,MD04Mi,MD04Ni); n=3,1

ME04

60. Relationship themes w/friends: f1 = narcissism - pres
ME04f1p = mean(ME04Bp,ME04Dp,ME04Hp,ME04Ip); n=4,NA
61. Relationship themes w/friends: f2 = belonging to peer group - pres
ME04f2p = mean(ME04Lp,ME04Mp,ME04Np); n=3,NA
62. Relationship themes w/friends: f3 = threatened - pres
ME04f3p = mean(ME04Ap,ME04Cp,ME04Kp); n=3,NA
63. Relationship themes w/friends: f4 = separateness - pres
ME04f4p = mean(ME04Ep,ME04Fp,ME04Gp,ME04Jp); n=4,NA
64. Relationship themes w/friends: f1 = narcissism - int
ME04f1i = mean(ME04Bi,ME04Di,ME04Hi,ME04Ii); n=4,NA
65. Relationship themes w/friends: f2 = belonging to peer group - int
ME04f2i = mean(ME04Li,ME04Mi,ME04Ni); n=3,NA
66. Relationship themes w/friends: f3 = threatened - int
ME04f3i = mean(ME04Ai,ME04Ci,ME04Ki); n=3,NA
67. Relationship themes w/friends: f4 = separateness - int
ME04f4i = mean(ME04Ei,ME04Fi,ME04Gi,ME04Ji); n=4,NA

MF04

68. Relationship themes w/sexual relations: f1=fear of intimacy - pres
MF04f1p = mean(MF04Cp,MF04Fp,MF04Gp,MF04Jp,MF04Kp,MF04Np); n=6,NA
69. Relationship themes w/sexual relations: f2=fear of being unloved - pres
MF04f2p = mean(MF04Bp,MF04Ep,MF04Hp,MF04Ip); n=4,NA
70. Relationship themes w/sexual relations: f3=fear of attack - pres
MF04f3p = mean(MF04Ap,MF04Dp); n=2,NA
71. Relationship themes w/sexual relations: f4=wish for stability - pres
MF04f4p = mean(MF04Lp,MF04Mp); n=2,NA
72. Relationship themes w/sexual relations: f1=fear of intimacy - int
MF04f1i = mean(MF04Ci,MF04Fi,MF04Gi,MF04Ji,MF04Ki,MF04Ni); n=6,NA
73. Relationship themes w/sexual relations: f2=fear of being unloved - int
MF04f2i = mean(MF04Bi,MF04Ei,MF04Hi,MF04Ii); n=4,NA
74. Relationship themes w/sexual relations: f3=fear of attack - int
MF04f3i = mean(MF04Ai,MF04Di); n=2,NA
75. Relationship themes w/sexual relations: f4=wish for stability - int
MF04f4i = mean(MF04Li,MF04Mi); n=2,NA

MG

76. Problems w/money and work - pres
MGp = mean(MG02p,MG03p,MG04p,MG07p,MG08p); n=5,NA
77. Problems w/money and work - int
MGi = mean(MG02i,MG03i,MG04i,MG07i,MG08i); n=5,NA

MH03 = Seriousness of life event/situation

(1=psychologically threatening,2=quite serious,3=very serious,4=extremely serious)

78. Life event: content weighted by seriousness - pres
MH03ap = sum(MH03Ap*4,MH03Bp*3,MH03Cp*2,MH03Dp)
79. Life event: seriousness w/max score - pres
MH03bp = seriousness item with maximum pres score
80. Capacity to cope (3=realist, 2=unrealist, 1=ineffect) - pres
MH04p = capacity to cope with maximum pres score
81. Life event: content weighted by seriousness - int
MH03ai = sum(MH03Ai*4,MH03Bi*3,MH03Ci*2,MH03Di)
82. Life event: seriousness w/max score - int
MH03bi = seriousness item with maximum int score
83. Capacity to cope (3=realist, 2=unrealist, 1=ineffect) - int
MH04i = capacity to cope with maximum int score

MH05

84. Affective reaction to event: f1=full of affect - pres
MH05f1p = mean(MH05Bp,MH05Ep,3-MH05Fp); n=3
85. Affective reaction to event: f2=depressed/suicidal - pres
MH05f2p = mean(MH05Cp,MH05Dp); n=2
86. Affective reaction to event: f3=not happy - pres
MH05f3p = 3-MH05Ap
87. Affective reaction to event: f1=full of affect - int
MH05f1i = mean(MH05Bi,MH05Ei,3-MH05Fi); n=3
88. Affective reaction to event: f2=depressed/suicidal - int
MH05f2i = mean(MH05Ci,MH05Di); n=2
89. Affective reaction to event: f3=not happy - int
MH05f3i = 3-MH05Ai

MI02

90. Gender, age, race: f1=patient's culture and race - pres
MI02f1p = mean(MI02Ep,MI02Gp); n=2,NA
91. Gender, age, race: f2=analyst's culture and race - pres
MI02f2p = mean(MI02Dp,MI02Fp); n=2,NA
92. Gender, age, race: f2=gender - pres
MI02f3p = mean(MI02Ap,MI02Bp); n=2,NA
93. Gender, age, race: f2=age - pres
MI02f4p = MI02Cp; n=1,NA
94. Gender, age, race: f1=patient's culture and race - int
MI02f1i = mean(MI02Ei,MI02Gi); n=2,NA
95. Gender, age, race: f2=analyst's culture and race - int
MI02f2i = mean(MI02Di,MI02Fi); n=2,NA
96. Gender, age, race: f2=gender - int
MI02f3i = mean(MI02Ai,MI02Bi); n=2,NA
97. Gender, age, race: f2=age - int
MI02f4i = MI02Ci; n=1,NA

MJ

98. Content for adult identity - pres
MJ02p = mean(MJ02Ap,MJ02Bp,MJ02Cp,MJ02Dp,MJ02Ep); n=5,NA
99. Identity shift, weight by: 0=none,1=defens,2=pos but prim,3=mature - pres
MJ03p = mean(MJ03Ap*3,MJ03Bp*2,MJ03Cp*1); n=6,NA
100. Content for adult identity - int
MJ02i = mean(MJ02Ai,MJ02Bi,MJ02Ci,MJ02Di,MJ02Ei); n=5,NA
101. Identity shift, weight by: 0=none,1=defens,2=pos but prim,3=mature - int
MJ03i = mean(MJ03Ai*3,MJ03Bi*2,MJ03Ci*1); n=6,NA

MK1

102. Sexual content (-1=avoided, 0=none, 1=implied, 2=explicit)
MK1 = -1 if MK01 is yes, 1 if MK04 is yes, 2 if MK03 is yes

MK05

103. Sexual content: f1 = perverse - pres
MK05f1p = mean(MK05J1p,MK05J2p,MK05J3p,MK05J4p,MK05J5p); n=5,4
104. Sexual content: f2 = sexual fantasy/homosexuality - pres
MK05f2p = mean(MK05Ap,MK05Ep,MK05Fp,MK05Hp); n=4,2
105. Sexual content: f3 = sexual life - pres
MK05f3p = mean(MK05Bp,MK05Dp); n=2
106. Sexual content: f1 = perverse - int
MK05f1i = mean(MK05J1i,MK05J2i,MK05J3i,MK05J4i,MK05J5i); n=5,4
107. Sexual content: f2 = sexual fantasy/homosexuality - int
MK05f2i = mean(MK05Ai,MK05Ei,MK05Fi,MK05Hi); n=4,2
108. Sexual content: f3 = sexual life - int
MK05f3i = mean(MK05Bi,MK05Di); n=2

MK06

109. Motivation for sexual material: f1 = inhibition/anxiety - pres
MK06f1p = mean(MK06Ap,MK06Bp,MK06Cp); n=3
110. Motivation for sexual material: f2 = shame/ambivalence - pres
MK06f2p = mean(MK06Dp,MK06Ep); n=2
111. Motivation for sexual material: f1 = inhibition/anxiety - int
MK06f1i = mean(MK06Ai,MK06Bi,MK06Ci); n=3
112. Motivation for sexual material: f2 = shame/ambivalence - int
MK06f2i = mean(MK06Di,MK06Ei); n=2

ML01

113. Discussion of tx param - non-discussion - pres
ML01p = mean(ML01Ap,3-ML01Bp)
114. Discussion of tx param - non-discussion - int
ML01i = mean(ML01Ai,3-ML01Bi)

ML03

115. Discussion of tx param, assoc affect: f1 = sad/angry - pres
ML03f1p = mean(ML03Bp,ML03Cp,ML03Dp,3-ML03Hp); n=4,2
116. Discussion of tx param, assoc affect: f2 = guilty/anxious/confused - pres
ML03f2p = mean(ML03Ep,ML03Fp,ML03Gp); n=3,1
117. Discussion of tx param, assoc affect: f1 = sad/angry - int
ML03f1i = mean(ML03Bi,ML03Ci,ML03Di,3-ML03Hi); n=4,2
118. Discussion of tx param, assoc affect: f2 = guilty/anxious/confused - int
ML03f2i = mean(ML03Ei,ML03Fi,ML03Gi); n=3,1

Preconscious Content

PA01

119. General transference
PA01 = mean(PA01Ap,PA01Bp,PA01Cp,PA02,PA03,PA04,PA05,PA06); n=5,4.33

PA02

120. Positive wishes towards analyst: f1 = loving/erotic - pres
PA02f1p = mean(PA02Ap,PA02Bp,PA02Cp,PA02Dp,PA02Hp); n=5
121. Positive wishes towards analyst: f2 = idealization/identification - pres
PA02f2p = mean(PA02Ep,PA02Gp); n=2
122. Positive wishes towards analyst: f3 = dependence - pres
PA02f3p = PA02Fp; n=1,NA
123. Positive wishes towards analyst: f1 = loving/erotic - int
PA02f1i = mean(PA02Ai,PA02Bi,PA02Ci,PA02Di,PA02Hi); n=5
124. Positive wishes towards analyst: f2 = idealization/identification - int
PA02f2i = mean(PA02Ei,PA02Gi); n=2
125. Positive wishes towards analyst: f3 = dependence - int
PA02f3i = PA02Fi; n=1,NA

PA03

126. Transference with anxiety: f1 = paranoid - pres
PA03f1p = mean(PA03Ap,PA03Dp,PA03Ep,PA03Ip); n=4,3
127. Transference with anxiety: f2 = fear of rejection - pres
PA03f2p = mean(PA03Bp,PA03Fp,PA03Gp); n=3
128. Transference with anxiety: f3 = projected aggression - pres
PA03f3p = mean(PA03Cp,PA03Hp); n=2
129. Transference with anxiety: f1 = paranoid - int
PA03f1i = mean(PA03Ai,PA03Di,PA03Ei,PA03Ii); n=4,3
130. Transference with anxiety: f2 = fear of rejection - int
PA03f2i = mean(PA03Bi,PA03Fi,PA03Gi); n=3
131. Transference with anxiety: f3 = projected aggression - int
PA03f3i = mean(PA03Ci,PA03Hi); n=2

PA04

132. Transference with fighting: f1 = helpless analyst - pres
PA04f1p = mean(PA04Fp,PA04Gp); n=2
133. Transference with fighting: f2 = victory in rivalry - pres
PA04f2p = mean(PA04Ap,PA04Bp); n=2
134. Transference with fighting: f1 = helpless analyst - int
PA04f1i = mean(PA04Fi,PA04Gi); n=2
135. Transference with fighting: f2 = victory in rivalry - int
PA04f2i = mean(PA04Ai,PA04Bi); n=2

PA05

136. Trans w/resent: f1 = derogatory towards analyst - pres
PA05f1p = mean(PA05Dp,PA05Ep,PA05Fp); n=3
137. Trans w/resent: f2 = abandoned by analyst - pres
PA05f2p = mean(PA05Cp,PA05Ip); n=2
138. Trans w/resent: f3 = loss of contact - pres
PA05f3p = mean(PA05Bp,PA05Hp); n=2,1
139. Trans w/resent: f4 = dependence - pres
PA05f4p = PA05Ap; n=1,NA
140. Trans w/resent: f1 = derogatory towards analyst - int
PA05f1i = mean(PA05Di,PA05Ei,PA05Fi); n=3
141. Trans w/resent: f2 = abandoned by analyst - int
PA05f2i = mean(PA05Ci,PA05Ii); n=2
142. Trans w/resent: f3 = loss of contact - int
PA05f3i = mean(PA05Bi,PA05Hi); n=2,1
143. Trans w/resent: f4 = dependence - int
PA05f4i = PA05Ai; n=1,NA

PA06

144. Primitive transferences - pres
PA06p = mean(PA06Ap,PA06Bp,PA06Cp,PA06Dp,PA06Ep,PA06Fp); n=6
145. Primitive transferences - int
PA06i = mean(PA06Ai,PA06Bi,PA06Ci,PA06Di,PA06Ei,PA06Fi); n=6

PB01

146. Quality of transference shift - pres
PB01p = mean(PB01Cp,PB01Dp,-PB01Bp,-2*PB01Ap); n=2,NA
147. Quality of transference shift - int
PB01i = mean(PB01Ci,PB01Di,PB01Bi,PB01Ai); n=4,NA

PC01

148. Clear emotional stance
PC01 = PC01B

PC02

149. Emotional stance: sadness - pres
PC02p = mean(PC02Ap,PC02Bp,PC02Cp,PC02Dp); n=4
150. Emotional stance: sadness - int
PC02i = mean(PC02Ai,PC02Bi,PC02Ci,PC02Di); n=4

PC03

151. Primitive emotional stance: f1=fear/shame - pres
PC03f1p = mean(PC03Bp,PC03Cp,PC03Ip,PC03Jp,PC03Kp); n=5,2
152. Primitive emotional stance: f2=manic grandiosity - pres
PC03f2p = mean(PC03Dp,PC03Ep); n=2
153. Primitive emotional stance: f3=existential anxiety - pres
PC03f3p = mean(PC03Ap,PC03Gp,PC03Hp); n=3,1
154. Primitive emotional stance: f1=fear/shame - int
PC03f1i = mean(PC03Bi,PC03Ci,PC03Ii,PC03Ji,PC03Ki); n=5,2
155. Primitive emotional stance: f2=manic grandiosity - int
PC03f2i = mean(PC03Di,PC03Ei); n=2

156. Primitive emotional stance: f3=existential anxiety - int
PC03f3i = mean(PC03Ai,PC03Gi,PC03Hi); n=3,1

PC04

157. Emotional stance: anger - pres
PC04p = mean(PC04Ap,PC04Bp,3-PC04Cp,PC04Dp); n=4,NA

158. Emotional stance: anger - int
PC04i = mean(PC04Ai,PC04Bi,3-PC04Ci,PC04Di); n=4,NA

PD

159. Defences: f1=projection, reaction formation, denial - pres
PDF1p = mean(PD03p,PD05p,PD06p,PD11p,PD13p); n=5

160. Defences: f2=isolation - pres
PDF2p = mean(PD01p,PD12p); n=2,1

161. Defences: f3=splitting - pres
PDF3p = mean(PD09p,PD10p,PD15p); n=3,1

162. Defences: f4=regression, projective identification - pres
PDF4p = mean(PD04p,PD08p,PD14p); n=3

163. Defences: f5=externalisation - pres
PDF5p = PD02p

164. Defences: f1=projection, reaction formation, denial - int
PDF1i = mean(PD03i,PD05i,PD06i,PD11i,PD13i); n=5

165. Defences: f2=isolation - int
PDF2i = mean(PD01i,PD12i); n=2,1

166. Defences: f3=splitting - int
PDF3i = mean(PD09i,PD10i,PD15i); n=3,1

167. Defences: f4=regression, projective identification - int
PDF4i = mean(PD04i,PD08i,PD14i); n=3

168. Defences: f5=externalisation - int
PDF5i = PD02i

----- Unconscious Content -----

UA

169. UCS themes - object relations: f1=narcissistic/infantile - pres
Uaf1p = mean(UA06p,UA08p,UA09p); n=3,1

170. UCS themes - object relations: f2=valued object - pres
Uaf2p = mean(UA03p,UA05p,UA07p); n=3

171. UCS themes - object relations: f3=greed/envy - pres
Uaf3p = UA04p

172. UCS themes - object relations: f1=narcissistic/infantile - int
Uaf1i = mean(UA06i,UA08i,UA09i); n=3,1

173. UCS themes - object relations: f2=valued object - int
Uaf2i = mean(UA03i,UA05i,UA07i); n=3

174. UCS themes - object relations: f3=greed/envy - int
Uaf3i = UA04i

UB

175. UCS themes - reactions to aggression: f1=destructive - pres
UBf1p = mean(UB02p,UB03p,UB04p,UB08p,UB09p); n=5

176. UCS themes - reactions to aggression: f2=anxious - pres
UBf2p = mean(UB05p,UB06p,UB07p); n=3,NA

177. UCS themes - reactions to aggression: f3=masochistic - pres
UBf3p = mean(UB11p,UB12p); n=2

178. UCS themes - reactions to aggression: f4=depression - pres
UBf4p = UB10p

179. UCS themes - reactions to aggression: f1=destructive - int
UBf1i = mean(UB02i,UB03i,UB04i,UB08i,UB09i); n=5

180. UCS themes - reactions to aggression: f2=anxious - int
UBf2i = mean(UB05i,UB06i,UB07i); n=3,NA

181. UCS themes - reactions to aggression: f3=masochistic - int
 UBf3i = mean(UB11i,UB12i); n=2

182. UCS themes - reactions to aggression: f4=depression - int
 UBf4i = UB10i

UC

183. UCS themes - related to sexuality: f1=Oedipal - pres
 UCf1p = mean(UC02p,UC03p,UC05p,UC07p,UC09p,UC10p,UC11p); n=7,4

184. UCS themes - related to sexuality: f2=in support of identity - pres
 UCf2p = mean(UC08p,UC12p,UC13p); n=3,NA

185. UCS themes - related to sexuality: f3=masochism - pres
 UCf3p = mean(UC01p,UC06p); n=2

186. UCS themes - related to sexuality: f1=Oedipal - int
 UCf1i = mean(UC02i,UC03i,UC05i,UC07i,UC09i,UC10i,UC11i); n=7,4

187. UCS themes - related to sexuality: f2=in support of identity - int
 UCf2i = mean(UC08i,UC12i,UC13i); n=3,NA

188. UCS themes - related to sexuality: f3=masochism - int
 UCf3i = mean(UC01i,UC06i); n=2

UD

189. UCS themes - related to self & self-esteem: f1=low self-esteem - pres
 Udf1p = mean(UD06p,UD07p,UD08p); n=3

190. UCS themes - related to self & self-esteem: f2=narcissistic - pres
 Udf2p = mean(UD09p,UD10p); n=2,NA

191. UCS themes - related to self & self-esteem: f3=feels obliterated - pres
 Udf3p = mean(UD03p,UD04p,UD05p); n=3

192. UCS themes - related to self & self-esteem: f4=self-critical - pres
 Udf4p = mean(UD01p,UD02p); n=2

193. UCS themes - related to self & self-esteem: f1=low self-esteem - int
 Udf1i = mean(UD06i,UD07i,UD08i); n=3

194. UCS themes - related to self & self-esteem: f2=narcissistic - int
 Udf2i = mean(UD09i,UD10i); n=2,NA

195. UCS themes - related to self & self-esteem: f3=feels obliterated - int
 Udf3i = mean(UD03i,UD04i,UD05i); n=3

196. UCS themes - related to self & self-esteem: f4=self-critical - int
 Udf4i = mean(UD01i,UD02i); n=2

UE

197. UCS themes - related to body - pres
 UEp = mean(UE01p,UE02p,UE03p,UE04p,UE05p); n=5,NA

198. UCS themes - related to body - int
 UEi = mean(UE01p,UE02p,UE03p,UE04p,UE05p); n=5,NA

UF01

199. Reaction of patient to interpretation: clearly feels helped - pres
 UF01p = mean(UF01Ap,UF01Bp,UF01Cp,UF01Dp); n=4,3

200. Reaction of patient to interpretation: clearly feels helped - int
 UF01i = mean(UF01Ai,UF01Bi,UF01Ci,UF01Di); n=4,3

UF02

201. Reaction of patient to interpretation: reports not feeling helped - pres
 UF02p = sum(UF02Ap,UF02Bp*2,UF02Cp*3)/6

202. Reaction of patient to interpretation: reports not feeling helped - int
 UF02i = sum(UF02Ai,UF02Bi*2,UF02Ci*3)/6

UF03

203. Reaction of patient to interpretation: f1=analyst feels pt was helped - contained - pres
 UF03f1p = mean(UF03Ap,UF03Ep); n=3,NA

204. Reaction of patient to interpretation: f2=analyst feels pt was helped - confronted - pres
 UF03f2p = mean(UF03Bp,UF03Cp,UF03Dp); n=2,NA

205. Reaction of patient to interpretation: f1=analyst feels pt was helped - contained - int
 UF03f1i = mean(UF03Ai,UF03Ei); n=3,NA

206. Reaction of patient to interpretation: f2=analyst feels pt was helped - confronted - int
 UF03f2i = mean(UF03Bi,UF03Ci,UF03Di); n=2,NA

UF04

207. Reaction of patient to interpretation: analyst feels pt reacts negatively - pres
 UF04p = mean(UF04Ap,UF04Bp,UF04Cp,UF04Dp,UF04Ep,UF04Fp); n=6,NA

208. Reaction of patient to interpretation: analyst feels pt reacts negatively - int
 UF04i = mean(UF04Ai,UF04Bi,UF04Ci,UF04Di,UF04Ei,UF04Fi); n=6,NA

UG

209. Style of interventions: f1=interprets object relations
 UGf1 = mean(UG08,UG09,UG10,UG11,UG12,UG13); n=6

210. Style of interventions: f2=interprets thinking, perception, behaviour
 UGf2 = mean(UG05,UG06,UG07); n=3

211. Style of interventions: f3=supportive
 UGf3 = mean(UG02,UG03,UG14,UG18); n=4,2

212. Style of interventions: f4=interprets defences
 UGf4 = mean(UG04,UG15,UG16,UG17); n=4,3

UHov

213. Analyst's feelings: overall
 UHov = mean(UH01,UH02,UH03,UH04,UH05,UH06,UH07,UH08,UH09,UH10,UH11,UH12,UH13,UH14,UH15,UH16,UH17,UH18,UH19,UH20,UH21,UH22); n=22,18

UH

214. Analyst's feelings: f1=feels attacked
 UHf1 = mean(UH09,UH10,UH11,UH12,UH13,UH19,UH22); n=7,5

215. Analyst's feelings: f2=inadequate/confused
 UHf2 = mean(UH14,UH16,UH17,UH20,UH21); n=5,4

216. Analyst's feelings: f3=feels loving/empathic
 UHf3 = mean(UH05,UH06,UH07,UH08); n=4

217. Analyst's feelings: f4=bored/cut off
 UHf4 = mean(UH15,UH18); n=2

UI

218. Patient impact: f1=bullying
 UIf1 = mean(UI03,UI04,UI05); n=3,NA

219. Patient impact: f2=narcissistic
 UIf2 = mean(UI06,UI07,UI08,UI10); n=4,NA

220. Patient impact: f3=analyst rejected
 UIf3 = UI09; n=1,NA

Calculate quality scales

221. Quality of week: pt's general stance
 J1 = J1

222. Quality of week: qual of manifest material
 J2 = J2

223. Quality of week: qual of anal's understanding
 J3 = J3

224. Quality of week: qual of anal's interpretive interventions
 J4 = J4

225. Quality of week: qual of pt's response to interpretation
 J5 = J4

226. Quality of week: overall qual of week
 J6 = J6

227. Quality of week: average of qual scales
 Javg = mean(J1,J2,J3,J4,J5,J6); n=6

APPENDIX 4.3B. FORMULAE FOR CALCULATION OF SUBSECTION SUMMARY SCORES (SSC).

#GA

1. Resistance-p
GAop = mean((1-GA01)*3, GA02*3, GA03*3, GA04*3, GA05p, GA06p, GA07p, GA08p, GA09p, GA10p, GA11p, GA12p, GA13p, GA14p); n=14,13
2. Resistance-i
GAoi = mean(GA05i, GA06i, GA07i, GA08i, GA09i, GA10i, GA11i, GA12i, GA13i, GA14i); n=10

#GB

3. Negative attitude-p
GBop = mean(3-GB01p, GB02p, GB03p, GB04p, GB05p, GB06p); n=6,4
4. Negative attitude-i
GBoi = mean(GB01i, GB02i, GB03i, GB04i, GB05i, GB06i); n=6,4

#GC

5. Time keeping problems-p
GCop = mean(GC02p, GC03p, GC04p, GC05p, GC06p); n=5
6. Time keeping problems-i
GCoi = mean(GC02i, GC03i, GC04i, GC05i, GC06i); n=5

#GD

7. Missed sessions, relevance-p
GDop = mean(GD01*3, GD05p, GD06p, GD07p); n=4,NA
8. Missed sessions, relevance-i
GDoi = mean(GD05i, GD06i, GD07i); n=3,NA

#GE

9. Bad behaviour in session-p
GEop = mean(GE02p, GE03p, GE04p, GE05p, GE06p, GE07p, GE08p, GE09p, GE10p, 3*GE11); n=10,8
10. Bad behaviour in session-i
GEoi = mean(GE02i, GE03i, GE04i, GE05i, GE06i, GE07i, GE08i, GE09i, GE10i); n=9,7

#GF

11. Quality of material-p
GFop = mean(GF02p, GF03p, GF04p, GF05p, GF06p, GF07p, GF08p); n=7
12. Quality of material-i
GFoi = mean(GF02i, GF03i, GF04i, GF05i, GF06i, GF07i, GF08i); n=7

#GG

13. Aggression and sexuality-p
GGop = mean(GG02p, GG03p, GG04p, GG05p, GG06p, GG07p, GG08p, GG09p, GG10p); n=9,8
14. Aggression and sexuality-i
GGoi = mean(GG02i, GG03i, GG04i, GG05i, GG06i, GG07i, GG08i, GG09i, GG10i); n=9,8

#GH

15. Immaturity of mental functioning-p
GHop = mean((1-GH01)*3, GH02*3, GH03*3, GH04*3, GH05p, GH06p, GH07p, GH08p, GH09p, GH10p, GH11p, GH12p); n=12,9
16. Immaturity of mental functioning-i
GHoi = mean(GH05i, GH06i, GH07i, GH08i, GH09i, GH10i, GH11i, GH12i); n=8,6

#MA

17. Body-p
MAop = mean(3*MA03, 2*MA02, MA09p, MA10p, MA11p, MA12p, MA13p, MA14p, MA15p, MA16p, MA17p); n=10,7
18. Body-i
MAoi = mean(MA09i, MA10i, MA11i, MA12i, MA13i, MA14i, MA15i, MA16i, MA17i); n=9,6

#MB

19. Self-esteem-p
 MBop = mean(MB02p,MB03p,MB04p,MB05p,MB06p,MB07p,MB08p); n=7,4

20. Self-esteem-i
 MBoi = mean(MB02i,MB03i,MB04i,MB05i,MB06i,MB07i,MB08i); n=7,4

#MC

21. Historical material-p
 MCop =
 mean(MC02Ap,MC02Bp,MC02Cp,MC02Dp,MC03Ap,MC03Bp,MC03Cp,MC03Dp,MC03Ep,MC03Fp,MC03Gp,
 MC04Ap,MC04Bp,MC04Cp,MC04Dp,MC04Ep,MC04Fp,MC04Gp,MC04Hp,MC04Ip,MC04Jp,MC04Kp);
 n=22,14

22. Historical material-i
 MCoi =
 mean(MC02Ai,MC02Bi,MC02Ci,MC02Di,MC03Ai,MC03Bi,MC03Ci,MC03Di,MC03Ei,MC03Fi,MC03Gi,
 MC04Ai,MC04Bi,MC04Ci,MC04Di,MC04Ei,MC04Fi,MC04Gi,MC04Hi,MC04Ii,MC04Ji,MC04Ki);
 n=22,14

#MD

23. Relationship w/family-p
 MDop =
 mean(MD02Ap,MD02Bp,MD03Ap,MD03Bp,MD03Cp,MD03Dp,MD04Ap,MD04Bp,MD04Cp,MD04Dp,MD04Ep,
 MD04Fp,MD04Gp,MD04Hp,MD04Ip,MD04Jp,MD04Kp,MD04Lp,MD04Mp,MD04Np,MD04Op,MD04Pp,MD04Q
 p); n=23,16

24. Relationship w/family-i
 MDoi =
 mean(MD02Ai,MD02Bi,MD03Ai,MD03Bi,MD03Ci,MD03Di,MD04Ai,MD04Bi,MD04Ci,MD04Di,MD04Ei,
 MD04Fi,MD04Gi,MD04Hi,MD04Ii,MD04Ji,MD04Ki,MD04Li,MD04Mi,MD04Ni,MD04Oi,MD04Pi,MD04Q
 i); n=23,16

#ME

25. Relationship w/friends-p
 MEop =
 mean(ME02Ap,ME02Bp,ME03Ap,ME03Bp,ME03Cp,ME03Dp,ME04Ap,ME04Bp,ME04Cp,ME04Dp,ME04Ep,
 ME04Fp,ME04Gp,ME04Hp,ME04Ip,ME04Jp,ME04Kp,ME04Lp,ME04Mp,ME04Np,ME04Op); n=21,NA

26. Relationship w/friends-i
 MEoi =
 mean(ME02Ai,ME02Bi,ME03Ai,ME03Bi,ME03Ci,ME03Di,ME04Ai,ME04Bi,ME04Ci,ME04Di,ME04Ei,
 ME04Fi,ME04Gi,ME04Hi,ME04Ii,ME04Ji,ME04Ki,ME04Li,ME04Mi,ME04Ni,ME04Oi); n=21,NA

#MF

27. Sexual relations-p
 MFop =
 mean(MF02Ap,MF02Bp,MF03Ap,MF03Bp,MF04Ap,MF04Bp,MF04Cp,MF04Dp,MF04Ep,MF04Fp,MF04Gp,
 MF04Hp,MF04Ip,MF04Jp,MF04Kp,MF04Lp,MF04Mp,MF04Np,MF04Op); n=19,NA

28. Sexual relations-i
 MFoi =
 mean(MF02Ai,MF02Bi,MF03Ai,MF03Bi,MF04Ai,MF04Bi,MF04Ci,MF04Di,MF04Ei,MF04Fi,MF04Gi,
 MF04Hi,MF04Ii,MF04Ji,MF04Ki,MF04Li,MF04Mi,MF04Ni,MF04Oi); n=19,NA

#MG

29. Money and work-p
 MGop = mean(MG02p,MG03p,MG04p,MG05p,MG06p,MG07p,MG08p); n=5,NA

30. Money and work-i
 MGoi = mean(MG02i,MG03i,MG04i,MG05i,MG06i,MG07i,MG08i); n=5,NA

#MH

31. Current life events-p
 MHop =
 mean(MH02p,MH03Ap,MH03Bp,MH03Cp,MH03Dp,MH03Ep,MH04Ap,MH04Bp,MH04Cp,MH04Dp,MH05Ap,
 MH05Bp,MH05Cp,MH05Dp,MH05Ep,MH05Fp); n=16,14

32. Current life events-i
 MHoi =
 mean(MH02i,MH03Ai,MH03Bi,MH03Ci,MH03Di,MH03Ei,MH04Ai,MH04Bi,MH04Ci,MH04Di,MH05Ai,
 MH05Bi,MH05Ci,MH05Di,MH05Ei,MH05Fi); n=16,14

#MI

33. Gender, age, and race-p
 MIop = mean(MI02Ap,MI02Bp,MI02Cp,MI02Dp,MI02Ep,MI02Fp,MI02Gp); n=7,NA

34. Gender, age, and race-i
 MIoi = mean(MI02Ai,MI02Bi,MI02Ci,MI02Di,MI02Ei,MI02Fi,MI02Gi); n=7,NA

#MJ

35. Adult identity-p
 MJop = mean(MJ02Ap,MJ02Bp,MJ02Cp,MJ02Dp,MJ02Ep,MJ03Ap,MJ03Bp,MJ03Cp,MJ03Dp);
 n=9,NA

36. Adult identity-i
 MJoi = mean(MJ02Ai,MJ02Bi,MJ02Ci,MJ02Di,MJ02Ei,MJ03Ai,MJ03Bi,MJ03Ci,MJ03Di);
 n=9,NA

#MK

37. Sexuality-p
 MKop =
 mean(MK01*3,MK03*3,MK04*3,MK05Ap,MK05Bp,MK05Cp,MK05Dp,MK05Ep,MK05Fp,MK05Gp,MK05Hp,
 MK05Ip,MK05J1p,MK05J2p,MK05J3p,MK05J4p,MK05J5p,MK05Kp,MK06Ap,MK06Bp,MK06Cp,MK06Dp,
 MK06Ep, MK06Fp); n=24,17

38. Sexuality-i
 MKoi =
 mean(MK05Ai,MK05Bi,MK05Ci,MK05Di,MK05Ei,MK05Fi,MK05Gi,MK05Hi,MK05Ii,MK05J1i,
 MK05J2i,MK05J3i,MK05J4i,MK05J5i,MK05Ki,MK06Ai,MK06Bi,MK06Ci,MK06Di,MK06Ei,MK06Fi);
 n=21,14

#ML

39. Treatment parameters-p
 MLOp =
 mean(ML01Ap,ML01Bp,ML02Ap,ML02Bp,ML02Cp,ML02Dp,ML02Ep,ML02Fp,ML03Ap,ML03Bp,ML03Cp,
 ML03Dp,ML03Ep,ML03Fp,ML03Gp,ML03Hp,ML03Ip); n=17,13

40. Treatment parameters-i
 MLoi =
 mean(ML01Ai,ML01Bi,ML02Ai,ML02Bi,ML02Ci,ML02Di,ML02Ei,ML02Fi,ML03Ai,ML03Bi,ML03Ci,
 ML03Di,ML03Ei,ML03Fi,ML03Gi,ML03Hi,ML03Ii); n=17,13

#PA1

41. General transference-p
 PA1op = mean(PA01Ap,PA01Bp,PA01Cp); n=3

42. General transference-i
 PA1oi = mean(PA01Ai,PA01Bi,PA01Ci); n=3

#PA2

43. Positive wishes towards analyst-p
 PA2op = mean(PA02Ap,PA02Bp,PA02Cp,PA02Dp,PA02Ep,PA02Fp,PA02Gp,PA02Hp); n=8,7

44. Positive wishes towards analyst-i
 PA2oi = mean(PA02Ai,PA02Bi,PA02Ci,PA02Di,PA02Ei,PA02Fi,PA02Gi,PA02Hi); n=8,7

#PA3

45. Transference w/anxiety-p
 PA3op = mean(PA03Ap,PA03Bp,PA03Cp,PA03Dp,PA03Ep,PA03Fp,PA03Gp,PA03Hp,PA03Ip);
 n=9,8

46. Transference w/anxiety-i
 PA3oi = mean(PA03Ai,PA03Bi,PA03Ci,PA03Di,PA03Ei,PA03Fi,PA03Gi,PA03Hi,PA03Ii);
 n=9,8

#PA4

47. Transference w/competition and aggression-p
 PA4op = mean(PA04Ap,PA04Bp,PA04Cp,PA04Dp,PA04Ep,PA04Fp,PA04Gp); n=7,5

48. Transference w/competition and aggression-i
 PA4oi = mean(PA04Ai,PA04Bi,PA04Ci,PA04Di,PA04Ei,PA04Fi,PA04Gi); n=7,5

#PA5

49. Transference w/resentment-p
 PA5op = mean(PA05Ap,PA05Bp,PA05Cp,PA05Dp,PA05Ep,PA05Fp,PA05Gp,PA05Hp,PA05Ip);
 n=9,7

50. Transference w/resentment-i
PA5oi = mean(PA05Ai, PA05Bi, PA05Ci, PA05Di, PA05Ei, PA05Fi, PA05Gi, PA05Hi, PA05Ii);
n=9,7

#PA6

51. Primitive transference-p
PA6op = mean(PA06Ap, PA06Bp, PA06Cp, PA06Dp, PA06Ep, PA06Fp); n=6

52. Primitive transference-i
PA6oi = mean(PA06Ai, PA06Bi, PA06Ci, PA06Di, PA06Ei, PA06Fi); n=6

#PB

53. Change in transference-p
PBop = mean(PB01Ap, PB01Bp, PB01Cp, PB01Dp); n=4, NA

54. Change in transference-i
PBoi = mean(PB01Ai, PB01Bi, PB01Ci, PB01Di); n=4, NA

#PC1

55. Affect
PC1o_mean((1-PC01A)*3, PC01B*3)

#PC2

56. Sadness-p
PC2op = mean(PC02Ap, PC02Bp, PC02Cp, PC02Dp, PC02Ep); n=5

57. Sadness-i
PC2oi = mean(PC02Ai, PC02Bi, PC02Ci, PC02Di, PC02Ei); n=5

#PC3

58. Primitive emotional stance-p
PC3op =
mean(PC03Ap, PC03Bp, PC03Cp, PC03Dp, PC03Ep, PC03Fp, PC03Gp, PC03Hp, PC03Ip, PC03Jp, PC03Kp)
; n=11,5

59. Primitive emotional stance-i
PC3oi =
mean(PC03Ai, PC03Bi, PC03Ci, PC03Di, PC03Ei, PC03Fi, PC03Gi, PC03Hi, PC03Ii, PC03Ji, PC03Ki)
; n=11,5

#PC4

60. Anger-p
PC4op = mean(PC04Ap, PC04Bp, PC04Cp, PC04Dp); n=4, NA

61. Anger-i
PC4oi = mean(PC04Ai, PC04Bi, PC04Ci, PC04Di); n=4, NA

#PDF1-f5: Defences

62. Defences: f1=projection, reaction formation, denial - pres
PDF1p = mean(PD03p, PD05p, PD06p, PD11p, PD13p); n=5

63. Defences: f2=isolation - pres
PDF2p = mean(PD01p, PD12p); n=2,1

64. Defences: f3=splitting - pres
PDF3p = mean(PD09p, PD10p, PD15p); n=3,1

65. Defences: f4=regression, projective identification - pres
PDF4p = mean(PD04p, PD08p, PD14p); n=3

66. Defences: f5=externalisation - pres
PDF5p = PD02p

67. Defences: f1=projection, reaction formation, denial - int
PDF1i = mean(PD03i, PD05i, PD06i, PD11i, PD13i); n=5

68. Defences: f2=isolation - int
PDF2i = mean(PD01i, PD12i); n=2,1

69. Defences: f3=splitting - int
PDF3i = mean(PD09i, PD10i, PD15i); n=3,1

70. Defences: f4=regression, projective identification - int
PDF4i = mean(PD04i, PD08i, PD14i); n=3

71. Defences: f5=externalisation - int
PDF5i = PD02i

```

#UA
72. UCS - general-p
    UAop = mean(UA01*3,UA02*3,UA03p,UA04p,UA05p,UA06p,UA07p,UA08p,UA09p); n=9,6
73. UCS - general-i
    UAoi = mean(UA03i,UA04i,UA05i,UA06i,UA07i,UA08i,UA09i); n=7,4

#UB
74. UCS - aggression-p
    UBop =
    mean(UB01p,UB02p,UB03p,UB04p,UB05p,UB06p,UB07p,UB08p,UB09p,UB10p,UB11p,UB12p);
    n=12,6
75. UCS - aggression-i
    UBoi =
    mean(UB01i,UB02i,UB03i,UB04i,UB05i,UB06i,UB07i,UB08i,UB09i,UB10i,UB11i,UB12i);
    n=12,6

#UC
76. UCS - sexuality-p
    UCop =
    mean(UC01p,UC02p,UC03p,UC04p,UC05p,UC06p,UC07p,UC08p,UC09p,UC10p,UC11p,UC12p,UC13p
    ); n=13,7
77. UCS - sexuality-i
    UCoi =
    mean(UC01i,UC02i,UC03i,UC04i,UC05i,UC06i,UC07i,UC08i,UC09i,UC10i,UC11i,UC12i,UC13i
    ); n=13,7

#UD
78. UCS - self and self-esteem-p
    UDop = mean(UD01p,UD02p,UD03p,UD04p,UD05p,UD06p,UD07p,UD08p,UD09p,UD10p); n=10,6
79. UCS - self and self-esteem-i
    UDoi = mean(UD01i,UD02i,UD03i,UD04i,UD05i,UD06i,UD07i,UD08i,UD09i,UD10i); n=10,6

#UE
80. UCS - body-p
    UEop = mean(UE01p,UE02p,UE03p,UE04p,UE05p); n=5,NA
81. UCS - body-i
    UEoi = mean(UE01i,UE02i,UE03i,UE04i,UE05i); n=5,NA

#UF1
82. Pt feels helped-p
    UF1p = mean(UF01Ap,UF01Bp,UF01Cp,UF01Dp); n=4,3
83. Pt feels helped-i
    UF1i = mean(UF01Ai,UF01Bi,UF01Ci,UF01Di); n=4,3

#UF2
84. Pt reports not feeling helped-p
    UF2p = mean(UF02Ap,UF02Bp,UF02Cp); n=3
85. Pt reports not feeling helped-i
    UF2i = mean(UF02Ai,UF02Bi,UF02Ci); n=3

#UF3
86. An feels patient was helped-p
    UF3p = mean(UF03Ap,UF03Bp,UF03Cp,UF03Dp,UF03Ep,UF03Fp); n=6,NA
87. An feels patient was helped-i
    UF3i = mean(UF03Ai,UF03Bi,UF03Ci,UF03Di,UF03Ei,UF03Fi); n=6,NA

#UF4
88. An feels patient reacts negatively-p
    UF4p = mean(UF04Ap,UF04Bp,UF04Cp,UF04Dp,UF04Ep,UF04Fp,UF04Gp); n=7
89. An feels patient reacts negatively-i
    UF4i = mean(UF04Ai,UF04Bi,UF04Ci,UF04Di,UF04Ei,UF04Fi,UF04Gi); n=7

#UGf1-f4: styles of intervention

```

90. Style of interventions: f1=interprets object relations
 UGf1 = (mean(UG08,UG09,UG10,UG11,UG12,UG13)-1)*0.75; n=6

91. Style of interventions: f2=interprets thinking, perception, behaviour
 UGf2 = (mean(UG05,UG06,UG07)-1)*0.75; n=3

92. Style of interventions: f3=supportive
 UGf3 = (mean(UG02,UG03,UG14,UG18)-1)*0.75; n=4,2

93. Style of interventions: f4=interprets defences
 UGf4 = (mean(UG04,UG15,UG16,UG17)-1)*0.75; n=4,3

#UHf1-f4: types of countertransference

94. Countertransference: attacked,disgusted
 UHf1o = (mean(UH09,UH10,UH11,UH12,UH13,UH19,UH22)-1)*0.75; n=7,5

95. Countertransference: inadequate/confused
 UHf2o = (mean(UH14,UH16,UH17,UH20,UH21)-1)*0.75; n=5,4

96. Countertransference: loving/empathic
 UHf3o = (mean(UH05,UH06,UH07,UH08)-1)*0.75; n=4

97. Countertransference: bored/cutoff
 UHf4o = (mean(UH15,UH18)-1)*0.75; n=2

#UIf1-f3: types of patient impact

98. Bullying
 UIf1o = (mean(UI03,UI04,UI05)-1)*0.75; n=3,NA

99. Narcissistic
 UIf2o = (mean(UI06,UI07,UI08,UI10)-1)*0.75; n=4,NA

100. Analyst rejected
 UIf3o = (UI09-1)*0.75

#J

101. Quality of week: average of qual scales
 Javg = (mean(J1,J2,J3,J4,J5,J6)-1)*0.75; n=6

APPENDIX 4.3C. FORMULAE FOR CALCULATION OF GLOBAL FACTORS BASED ON SUBSECTION SUMMARY SCORES (SOF).

1. Resistance

$$\text{SOFN1} = 10 * \text{mean}(\text{GAop}, \text{GBop}, \text{GOp}, \text{GHop}, \text{PA4op}, \text{PA5op}, \text{PA6op}, \text{PDF4p}, \text{UF2p}, \text{UF4p}, \text{UHf1o}, \text{UHf2o}, \text{UHf4o}, \text{UIf1o}, \text{UIf2o}, \text{UIf3o}, 3\text{-GFop}, 3\text{-MAop}, 3\text{-MDop}, 3\text{-MEop}, 3\text{-MJop}, 3\text{-PA2op}, 3\text{-PC1op}, 3\text{-PC2op}, 3\text{-UF1p}, 3\text{-UF3p}, 3\text{-UGf1}, 3\text{-UGf3}, 3\text{-UHf3o}); n=29$$

2. Coherent Ucs Material

$$\text{SOFN2} = 10 * \text{mean}(\text{MBop}, \text{MIop}, \text{MKop}, \text{PA3op}, \text{PC3op}, \text{PC4op}, \text{UAop}, \text{UBop}, \text{UCop}, \text{UDop}, \text{UEop}, 3\text{-PDF2p}, 3\text{-PDF5p}, 3\text{-UGf2}, 3\text{-UGf4}); n=15$$

3. Resistance / Coherent UCS Material

$$\text{SOFNrat} = \text{SOFN1} / \text{SOFN2}$$

APPENDIX 4.3D. FORMULAE FOR CALCULATION OF SCALES FOR HYPOTHESIS TESTING.

1. Therapist: Jones 'Dynamic technique'

$$\text{TxJones} = \text{mean}(\text{mean}(\text{PC2oi}, \text{PC3oi}, \text{PC4oi}), \text{mean}(\text{UAoi}, \text{UBoi}, \text{UCoi}, \text{UDoi}, \text{UEoi}), \text{sumv1}(0.75 * (\text{UG15-1}) * 0.75, (\text{UG16-1}) * 0.75, (\text{UG17-1}) * 0.75, \text{Pdf1i}, \text{Pdf2i}, \text{Pdf3i}, \text{Pdf4i}, \text{Pdf5i}), \text{mean}(\text{MCoi}, (\text{UG12-1}) * 0.75), \text{mean}(\text{PA1oi}, \text{PA2oi}, \text{PA3oi}, \text{PA4oi}, \text{PA5oi}, \text{PA6oi}, \text{PBoi}, (\text{UG09-1}) * 0.75, (\text{UG10-1}) * 0.75, (\text{UG11-1}) * 0.75), \text{mean}(\text{GAoi}, \text{GBoi}, \text{GCoi}, \text{GDoi}, \text{GEoi}, \text{GGoi}), \text{mean}(\text{GF02i}, (\text{J3-1}) * 0.75, (\text{UG13-1}) * 0.75), \text{mean}(\text{GF05i}, \text{GF06i}, \text{GF07i}, \text{GF08i}), \text{mean}(\text{MFOi}, \text{MKoi}, \text{UCoi}), \text{mean}(\text{MAoi}, \text{MBoi}, \text{UDoi}, \text{UEoi}), \text{mean}(\text{MC04Ci}, \text{MD04Bi}, \text{MD04Hi}, \text{ME04Ci}, \text{ME04Ji}, \text{MF04Di}, \text{MF04Ji}, \text{MK06Di}, \text{PC02Ci}, \text{PC02Di}, \text{PC03Fi}, \text{PC03Ii}, \text{PC03Ji}, \text{UF04Bi}); n=11$$
2. Therapist: general interpretation

$$\text{TxInt} = \text{mean}(\text{GAoi}, \text{GBoi}, \text{GCoi}, \text{GDoi}, \text{GEoi}, \text{GFOi}, \text{GGoi}, \text{GHoi}, \text{MAoi}, \text{MBoi}, \text{MCoi}, \text{MDoi}, \text{MEoi}, \text{MFOi}, \text{MGoi}, \text{MHOi}, \text{MIOi}, \text{MJoi}, \text{MKoi}, \text{MLOi}, \text{PA1oi}, \text{PA2oi}, \text{PA3oi}, \text{PA4oi}, \text{PA5oi}, \text{PA6oi}, \text{PBoi}, \text{PC2oi}, \text{PC3oi}, \text{PC4oi}, \text{Pdf1i}, \text{Pdf2i}, \text{Pdf3i}, \text{Pdf4i}, \text{Pdf5i}, \text{UAoi}, \text{UBoi}, \text{UCoi}, \text{UDoi}, \text{UEoi}, \text{UF1i}, \text{UF2i}, \text{UF3i}, \text{UF4i}, \text{UGf1}, \text{UGf2}, \text{UGf3}, \text{UGf4}); n=48$$
3. Therapist: transference interpretation

$$\text{TxTrans} = \text{mean}(\text{PA1oi}, \text{PA2oi}, \text{PA3oi}, \text{PA4oi}, \text{PA5oi}, \text{PA6oi}, \text{PBoi}, (\text{UG09-1}) * 0.75, (\text{UG10-1}) * 0.75, (\text{UG11-1}) * 0.75); n=10$$
4. Therapist: relationship interpretation

$$\text{TxRelat} = \text{mean}(\text{MDoi}, \text{MEoi}, \text{MFOi}, \text{MKoi}, (\text{UG13-1}) * 0.75, (\text{UG14-1}) * 0.75); n=6$$
5. Therapist: exploration

$$\text{TxExp} = \text{UGf2}$$
6. Therapist: supportive

$$\text{TxSup} = \text{UGf3}$$
7. Therapist: positive view of therapy

$$\text{TxPos} = \text{mean}(\text{UF3p}, \text{UHf3o}, \text{Javg}, 3\text{-UHf1o}, 3\text{-UHf2o}, 3\text{-UHf4o}, 3\text{-UIf1o}, 3\text{-UIf2o}, 3\text{-UIf3o}); n=9$$
8. Patient: Jones 'Dynamic material'

$$\text{PTJones} = \text{mean}(\text{mean}(\text{PC2op}, \text{PC3op}, \text{PC4op}), \text{mean}(\text{UAop}, \text{UBop}, \text{UCop}, \text{UDop}, \text{UEop}), \text{mean}(\text{Pdf1p}, \text{Pdf2p}, \text{Pdf3p}, \text{Pdf4p}, \text{Pdf5p}), \text{MCop}, \text{mean}(\text{PA1op}, \text{PA2op}, \text{PA3op}, \text{PA4op}, \text{PA5op}, \text{PA6op}, \text{PBop}), \text{mean}(\text{GAop}, \text{GBop}, \text{GCop}, \text{GDop}, \text{GEop}, \text{GGop}), \text{GF02p}, \text{mean}(\text{GF05p}, \text{GF06p}, \text{GF07p}, \text{GF08p}), \text{mean}(\text{MFOp}, \text{MKop}, \text{UCop}), \text{mean}(\text{MAop}, \text{MBop}, \text{UDop}, \text{UEop}), \text{mean}(\text{MC04Cp}, \text{MD04Bp}, \text{MD04Hp}, \text{ME04Cp}, \text{ME04Jp}, \text{MF04Dp}, \text{MF04Jp}, \text{MK06Dp}, \text{PC02Cp}, \text{PC02Dp}, \text{PC03Fp}, \text{PC03Ip}, \text{PC03Jp}, \text{UF04Bp}); n=11$$
9. Patient: transference themes

$$\text{PtTrans} = \text{mean}(\text{PA1op}, \text{PA2op}, \text{PA3op}, \text{PA4op}, \text{PA5op}, \text{PA6op}, \text{PBop}); n=7$$
10. Patient: relationships

$$\text{PtRelat} = \text{mean}(\text{MDop}, \text{MEop}, \text{MFOp}, \text{MKop}); n=4$$
11. Patient: regression

$$\text{PtReg} = \text{mean}(\text{GGop}, \text{GHop}); n=2$$
12. Patient: SOF Ratio Ucs/Regression to Resistance

$$\text{SOFNrat} = \text{SOFN1} / \text{SOFN2}$$
13. Patient: positive view of therapy

$$\text{PtPos} = \text{mean}(\text{PA2op}, \text{UF1p}, 3\text{-GBop}, 3\text{-PA3op}, 3\text{-PA4op}, 3\text{-PA5op}, 3\text{-PA6op}, 3\text{-UF2p}, 3\text{-UF4p}); n=9$$
14. Interaction: contract

$$\text{Cont} = \text{mean}(\text{MLOp}, \text{MLOi}); n=2$$
15. Interaction: alliance

$$\text{All} = (\text{J1-1}) * 0.75$$

APPENDIX 7.1. AAI CLASSIFICATION DATA

<i>Subject</i>	<i>Assessments</i>				
A	INITIAL U/F1a	FOLLOW 1 U/E2/E1	FOLLOW 2 U [?] /E2 ¹	FOLLOW 3 U [?] /E1/E2	TERM U [?] /E2
B	INITIAL U [?] /Ds3	TERM Interviewer errors			
C	INITIAL F2/F4	FOLLOW 1 U [?] /CC/E2/ Ds1/E1/Ds2	FOLLOW 2 U/CC/Ds1/ Ds2/E2/E1	FOLLOW 3 U/E1/E3	TERM E2
D	INITIAL U [?] /F/CC/E ¹	FOLLOW 1 U/F4b	FOLLOW 2 E2/E1	FOLLOW 3 U [?] /E1	TERM U/E3/E1/E2
E	INITIAL F5/F4b	FOLLOW 1 U/F2/Ds3	FOLLOW 2 U [?] /F ¹	FOLLOW 3 F4b	TERM U [?] /F4b
I	INITIAL U [?] /F2/F4b				
M	INITIAL Ds1/Ds3	TERM E3	PT FOLLOW 1 U ² /F3b/F1a		
O	INITIAL Recording failure	FOLLOW 1 U/E3/F4/F5	FOLLOW 2 U/E2/F5	FOLLOW 3 Recording failure	FOLLOW 4 F4b/F1a
P	INITIAL E1	FOLLOW 1 E2/E1	FOLLOW 2 E2/E1	TERM E2/E1	PT FOLLOW 1 F4/F3/F2
Q	INITIAL F3b	TERM F2	PT FOLLOW 1 E2	PT FOLLOW 2 Recording failure	
T	INITIAL Ds3/F1a	FOLLOW 1 CC/E1/ Ds1/E3	FOLLOW 2 CC ¹	TERM Ds2/Ds3	PT FOLLOW 1 Recording failure
U	INITIAL U/F4a/F4b	FOLLOW 1 U [?] /CC/ E2/Ds2	FOLLOW 2 F2	TERM U/F4	
V	INITIAL U [?] /F1b	PT FOLLOW 1 U/F2/F5			
X	INITIAL U [?] /F5/Ds3 ¹	FOLLOW 1 F2	TERM F2		

Table A7.1a. AAI classification by time for intensive patients.

¹ Tentative overall classifications due to interviewer errors

² Unresolved classification is due to loss that occurred since last assessment

<i>Subject</i>	<i>Assessments</i>			
F	INITIAL U?/E1	TERM U?/E1/E2	PT FOLLOW 1 U?/E3/ Ds3/CC	PT FOLLOW 2 E1
G	INITIAL U/CC/E1/ Ds4/Ds2			
H	INITIAL F4b	TERM F3b/F4b	PT FOLLOW 1 F4b	
J	INITIAL U?/E2/CC ¹			
K	INITIAL F2/Ds3	FOLLOW 1 E1/Ds2	FOLLOW 2 Ds1	
L	INITIAL U?/Ds3/F1a	TERM U/F4b	PT FOLLOW 1 F1b	
N	INITIAL U?/E3/E2 ¹	TERM U/E3a	PT FOLLOW 1 Recording failure	
R	INITIAL Ds3/F1b	TERM F4b	PT FOLLOW 1 F2/F4	
S	INITIAL F2/Ds3			
W	INITIAL U/CC/E3/ E1/Ds3/Ds2			
Y	INITIAL F2/Ds3	TERM CC/F2/Ds3		

Table A7.1b. AAI classification by time for non-intensive patients.

REFERENCES

- Ablon, S., & Jones, E. E. (1998). How expert clinicians' prototypes of an ideal treatment correlate with outcome in psychodynamic and cognitive-behavioral therapy. Psychotherapy Research, 8(1), 71-83.
- Ablon, S., & Jones, E. E. (1999). Psychotherapy process in the NIMH Collaborative Study. Journal of Consulting and Clinical Psychology, 67(1), 64-75.
- Achenbach, T. M., & McConaughy, S. H. (1997). Empirically Based Assessment of Child and Adolescent Psychopathology: Practical Applications (2nd ed.). Thousand Oaks, CA: Sage.
- Ackerman, S. J., & Hilsenroth, M. J. (2003). A review of therapist characteristics and techniques positively impacting the therapeutic alliance. Clinical Psychology Review, 23, 1-33.
- Adam, K. S., Sheldon-Keller, A. E., & West, M. (1996). Attachment organization and history of suicidal behavior in clinical adolescents. Journal of Consulting and Clinical Psychology, 64(2), 264-272.
- Adatto, C. (1980). Late adolescence to early adulthood. In S. Greenspan & G. Pollock (Eds.), Course of Life: Psychoanalytic Contributions Toward Understanding Personality Development (Vol. 2, pp. 463-474). Washington, D. C.: US Dept. Health & Human Services.
- Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (Eds.). (1978). Patterns of attachment: A psychological study of the Strange Situation. Hillsdale, NJ: Erlbaum.
- Alexander, F. (1937). Five-Year Report of the Chicago Institute for Psychoanalysis, 1932-1937. Chicago: Chicago Psychoanalytic Institute.
- Allen, J. P., Hauser, S. T., & Borman-Spurrell, E. (1996). Attachment theory as a framework for understanding sequelae of severe adolescent psychopathology: An 11-year follow-up study. Journal of Consulting and Clinical Psychology, 64(2), 254-263.
- Allison, G. H. (1994). On the homogenization of psychoanalysis and psychoanalytic psychotherapy: a review of some of the issues. Journal of the American Psychoanalytic Association, 42, 341-362.

- Anastasi, A., & Urbina, S. (1997). Psychological Testing (Seventh ed.). Upper Saddle River, NJ: Prentice Hall.
- Andrusyna, T. P., Tang, T. Z., DeRubeis, R. J., & Luborsky, L. (2001). The factor structure of the Working Alliance Inventory in cognitive-behavioral therapy. Journal of Psychotherapy Practice and Research, *10*(3), 173-178.
- Ankuta, G. Y., & Abeles, N. (1993). Client satisfaction, clinical significance, and meaningful change in psychotherapy. Professional Psychology: Research and Practice, *24*, 70-74.
- Anstadt, T., Merten, J., Ullrich, B., & Krause, R. (1997). Affective dyadic behavior, core conflictual relationship themes, and success of treatment. Psychotherapy Research, *7*(4), 397-417.
- APA. (1994). Diagnostic and Statistical Manual of Mental Disorders (Fourth ed.). Washington, DC: American Psychiatric Association.
- APA (Ed.). (2000). Handbook of Psychiatric Measures. Washington, DC: American Psychiatric Association.
- APA Task Force on Psychological Intervention Guidelines. (1995). Template for developing guidelines: Interventions for mental disorders and psychological aspects of physical disorders. Washington, DC: American Psychological Association.
- Appelbaum, A. (1994). Psychotherapeutic routes to structural change. Bulletin of the Menninger Clinic, *58*(1), 37-54.
- APsA. (1974). Psychoanalytic research. Homestead, Virginia: Conference on Psychoanalytic Education and Research, American Psychoanalytic Association.
- Bachelor, A. (1991). Comparison and relationship to outcome of diverse dimensions of the Helping Alliance as seen by client and therapist. Psychotherapy, *28*(4), 534-549.
- Bachrach, H. M. (1993). The Columbia Records Project and the evolution of psychoanalytic outcome research. Journal of the American Psychoanalytic Association, *41* supplement, 279-297.
- Bachrach, H. M., Galatzer-Levy, R., Skolnikoff, A., & Waldron, S., Jr. (1991). On the efficacy of psychoanalysis. Journal of the American Psychoanalytic Association, *39*(871-915).

- Bachrach, H. M., Weber, J. J., & Solomon, M. (1985). Factors associated with the outcome of psychoanalysis (clinical and methodological considerations): Report of the Columbia Psychoanalytic Center Research Project (IV). International Review of Psychoanalysis, *12*, 379-389.
- Baer, P., Dunbar, P. W., Hamilton, J. E. I., & Beutler, L. E. (1980). Therapists' perceptions of the psychotherapeutic process: Development of a psychotherapy process inventory. Psychological Reports, *46*, 563-570.
- Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (1993). A psychometric study of the Adult Attachment Interview: Reliability and discriminant validity. Developmental Psychology, *29*(5), 870-879.
- Bänninger-Huber, E. (1992). Prototypical Affective Microsequences in psychotherapeutic interaction. Psychotherapy Research, *2*(4), 291-306.
- Bänninger-Huber, E., & Widmer, C. (1997, June 19-21). Patterns of interaction and psychotherapeutic change. Paper presented at the 20th Ulm Workshop of Empirical Research in Psychoanalysis, Ulm, Germany.
- Barber, J. P., & Crits-Christoph, P. (1993). Advances in measures of psychodynamic formulations. Journal of Consulting and Clinical Psychology, *61*(4), 574-585.
- Barber, J. P., Foltz, C., & Weinryb, R. M. (1998). The Central Relationship Questionnaire: Initial report. Journal of Counseling Psychology, *45*(2), 131-142.
- Barlow, D. H., & Hersen, M. (1973). Single-case experiment designs: Uses in applied clinical research. Archives of General Psychiatry, *29*, 319-325.
- Baruch, G., Gerber, A., & Fearon, P. (1998). Adolescents who drop out of psychotherapy at a community-based psychotherapy centre: A preliminary investigation of the characteristics of early drop-outs, late drop-outs and those who continue treatment. British Journal of Medical Psychology, *71*(3), 233-245.
- Bateman, A., & Fonagy, P. (1999). Effectiveness of partial hospitalization in the treatment of borderline personality disorder - a randomised control trial. American Journal of Psychiatry, *156*(10), 1563-1569.

- Beck, A. T., & Steer, R. A. (1993). Beck Depression Inventory: Manual. San Antonio: The Psychological Corporation.
- Bellak, L., Hurvich, M., & Gediman, H. K. (1973). Ego functions in schizophrenics, neurotics, and normals. A systematic study of conceptual, diagnostic and therapeutic aspects. New York: John Wiley.
- Benjamin, L. S. (1974). Structural analysis of social behavior. Psychological Review, 81(5), 392-425.
- Benjamin, L. S. (1994). SASB: A bridge between personality theory and clinical psychology. Psychological Inquiry, 5(4), 273-316.
- Benjamin, L. S. (1996). Interpersonal Diagnosis and Treatment of Personality Disorders. New York: The Guilford Press.
- Benoit, D., & Parker, K. C. H. (1994). Stability and transmission of attachment across three generations. Child Development, 65, 1444-1456.
- Beres, D. (1968). Discussion of Engel 1968. Journal of the American Psychoanalytic Association, 16, 205-210.
- Berger, P., Sachs, G., Amering, M., Holzinger, A., Baknkier, B., & Katschnig, H. (2004). Personality disorder and social anxiety predict delayed response in drug and behavioral treatment of panic disorder. Journal of Affective Disorders, 80(1), 75-78.
- Berzins, J. I., Bednar, R. L., & Severy, L. J. (1975). The problem of intersources consensus in measuring therapeutic outcomes: New data and multivariate perspectives. Journal of Abnormal Psychology, 84(1), 10-19.
- Beutler, L. E. (1995). The germ theory myth and the myth of outcome homogeneity. Psychotherapy, 32(3), 489-494.
- Beutler, L. E., Malik, M., Alimohamed, S., Harwood, T. M., Talebo, H., Noble, S., & Wong, E. (2004). Therapist variables. In M. J. Lambert (Ed.), Bergin and Garfield's Handbook of psychotherapy and behavior change (Fifth ed., pp. 227-306). New York: John Wiley & Sons.

- Bieber, L., Dain, H. J., Dince, P. R., Drellich, M. G., Grand, H. G., Gundlach, R. H., Kremer, M. W., Rifkin, A. H., Wilbur, C. B., & Bieber, T. B. (1962). Homosexuality: A Psychoanalytic Study of Male Homosexuals. New York: Basic Books.
- Bifulco, A., Brown, G. W., & Harris, T. O. (1994). Childhood Experience of Care and Abuse (CECA): A retrospective interview measure. Journal of Child Psychology and Psychiatry, *35*(8), 1419-1435.
- Blatt, S. J. (1992). The differential effect of psychotherapy and psychoanalysis with anaclitic and introjective patients: The Menninger Psychotherapy Research Project revisited. Journal of the American Psychoanalytic Association, *40*, 691-724.
- Blatt, S. J. (2001). The effort to identify for Empirically Supported Psychological Treatments its implications for clinical research, practice, and training: Commentary on papers by Lester Luborsky and Hans H. Strupp. Psychoanalytic Dialogues, *11*(4), 635-646.
- Blatt, S. J., & Auerbach, J. S. (2003). Psychodynamic measures of therapeutic change. Psychoanalytic Inquiry, *23*(2), 268-307.
- Blatt, S. J., & Ford, R. Q. (1994). Therapeutic Change: An Object Relations Perspective. New York: Plenum Press.
- Blatt, S. J., Quinlan, D. M., Plikonis, P. A., & Shea, M. T. (1995). Impact of perfectionism and need for approval on the brief treatment of depression: The National Institute of Mental Health Treatment of Depression Collaborative Research Program revisited. Journal of Consulting and Clinical Psychology, *63*, 125-132.
- Blatt, S. J., & Shahar, G. (2004). Psychoanalysis - with whom, for what, and how? Comparisons with psychotherapy. Journal of the American Psychoanalytic Association, *52*(2), 393-447.
- Block, J. (1961/1978). The Q-sort Method in Personality Assessment and Psychiatric Research. Springfield, IL: Charles C. Thomas.
- Blos, P. (1967). The second individuation process of adolescence. Psychoanalytic Study of the Child, *22*, 162-186.

- Blos, P. (1977). When and how does adolescence end: Structural criteria for adolescent closure. In S. Feinstein & P. Giovacchini (Eds.), Adolescent Psychiatry (Vol. 5, pp. 5-17). New York: Jason Aronson.
- Blum, H. P. (1992). Psychic change: The analytic relationship(s) and agents of change. International Journal of Psycho-analysis, 73(255-264).
- Bøgwald, K.-P., Høgland, P., & Sørbye, Ø. (1999). Measurement of transference interpretations. Journal of Psychotherapy Practice and Research, 8(4), 264-273.
- Bond, M., Banon, E., & Grenier, M. (1998). Differential effects of interventions on the therapeutic alliance with patients with personality disorders. Journal of Psychotherapy Practice and Research, 7(4), 301-318.
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. Psychotherapy: Theory, Research, and Practice, 16(3), 252-260.
- Bowlby, J. (1960). Grief and mourning in infancy and early childhood. Psychoanalytic Study of the Child, 15, 9-52.
- Bowlby, J. (1969). Attachment. London: Hogarth Press.
- Bowlby, J. (1973). Attachment and loss: Vol. 2. Separation. New York: Basic Books.
- Bowlby, J. (1980). Attachment and loss: Vol. 3. Loss. New York: Basic Books.
- Bowlby, J. (1988). A Secure Base. New York: Basic Books.
- Brenman-Pick, I. (1985). Working through in the countertransference. International Journal of Psycho-Analysis, 66, 157-166.
- Bretherton, I. (1995). The origins of attachment theory: John Bowlby and Mary Ainsworth. In S. Goldberg & R. Muir & J. Kerr (Eds.), Attachment theory: Social, developmental, and clinical perspectives (pp. 45-84). Hillsdale, NJ: The Analytic Press.
- Brody, B. (1970). Freud's case load. Psychotherapy: Theory, Research, and Practice, 7, 8-12.
- Bryk, A. S., & Raudenbush, S. W. (1987). Application of hierarchical linear models to assessing change. Psychological Bulletin, 101(1), 147-158.

- Bryk, A. S., & Raudenbush, S. W. (2001). Hierarchical linear models: Applications and data analysis methods (Second ed.). Newbury Park, CA: Sage.
- Bryk, A. S., Raudenbush, S. W., & Congdon, R. T. (1996). HLM: Hierarchical linear and nonlinear modeling with the HLM/2L and HLM/3L programs. Chicago: Scientific Software International.
- Bucci, W. (1993). The development of emotional meaning in free association: A multiple code theory. In J. Gedo & A. Wilson (Eds.), Hierarchical Conceptions in Psychoanalysis (pp. 3-47). New York: Guilford Press.
- Bucci, W. (1997a). Patterns of discourse in "good" and troubled hours: A multiple code interpretation. Journal of the American Psychoanalytic Association, *45*, 155-187.
- Bucci, W. (1997b). Psychoanalysis and Cognitive Science: A Multiple Code Theory. New York: The Guilford Press.
- Bucci, W. (1998). Transformation of meanings in the analytic discourse. A strategy for research. Canadian Journal of Psychoanalysis, *6*, 233-260.
- Bucci, W. (2000). The need for a 'psychoanalytic psychology' in the cognitive science field. Psychoanalytic Psychology, *17*, 203-224.
- Bucci, W. (2001). Pathways of emotional communication. Psychoanalytic Inquiry, *20*, 40-70.
- Canfield, M. L., Walker, W. R., & Brown, L. G. (1991). Contingency interaction analysis in psychotherapy. Journal of Consulting and Clinical Psychology, *59*(1), 58-66.
- Carlson, E. A., & Sroufe, L. A. (1995). Contribution of attachment theory to developmental psychopathology. In D. Cicchetti & D. J. Cohen (Eds.), Developmental psychopathology: Vol 1. Theory and methods (pp. 581-617). New York: Wiley.
- Cartwright, R. D. (1966). A comparison of the response to psychoanalytic and client-centered psychotherapy. In L. A. Gottschalk & A. A. Auerbach (Eds.), Methods of research in psychotherapy (pp. 517-529). New York: Appleton-Century-Crofts.
- Cassidy, J. (1995). Attachment and generalized anxiety disorder. In D. Cicchetti & S. Toth (Eds.), Rochester Symposium on Developmental Psychopathology: Vol. 6. Emotion, cognition, and representation (pp. 343-370). Rochester, NY: University of Rochester Press.

- Castonguay, L. G., Goldfried, M. R., Wiser, S., Raue, P. J., & Hayes, A. M. (1996). Predicting the effect of cognitive therapy for depression: A study of unique and common factors. Journal of Consulting and Clinical Psychology, *64*(3), 497-504.
- Chambless, D. L., Sanderson, W. C., Shoham, V., Bennett Johnson, S., Pope, K. S., Crits-Christoph, P., Baker, M., Johnson, B., Woody, S. R., Sue, S., Beutler, L., Williams, D. A., & McCurry, S. (1996). An update on Empirically Validated Therapies. Clinical Psychologist, *49*, 5-18.
- Christensen, L., & Mendoza, J. L. (1986). A method of assessing change in a single subject: An alteration of the RC index [Letter to the editor]. Behavior Therapy, *17*, 305-308.
- Clyman, R. B. (1991). The procedural organization of emotions: A contribution from cognitive science to the psychoanalytic theory of therapeutic action. Journal of the American Psychoanalytic Association, *39*(supplement), 349-382.
- Cole-Detke, H., & Kobak, R. (1996). Attachment processes in eating disorder and depression. Journal of Consulting and Clinical Psychology, *64*(2), 282-290.
- Connolly, M. B. (1998). Therapist interventions in early session of brief supportive-expressive psychotherapy for depression. Journal of Psychotherapy Practice and Research, *7*(4), 290-300.
- Connolly, M. B., Crits-Christoph, P., Shappell, S., Barber, J. P., Luborsky, L., & Shaffer, C. (1999). Relation of transference interpretations to outcome in the early sessions of brief supportive-expressive psychotherapy. Psychotherapy Research, *9*(4), 485-495.
- Conway, M., & Ross, M. (1984). Getting what you want by revising what you had. Journal of Personality and Social Psychology, *47*, 738-748.
- Cook, T. D., & Campbell, D. T. (1979). Quasi-Experimentation: Design & Analysis Issues for Field Settings. Boston: Houghton Mifflin.
- Cooper, P., Osborn, M., Gath, D., & Feggetter, G. (1982). Evaluation of a modified self-report measure of social adjustment. British Journal of Psychiatry, *141*, 68-75.
- Coriat, I. H. (1917). Some statistical results of the psychoanalytic treatment of the psychoneuroses. Psychoanalytic Review, *4*, 209-216.

- Crits-Christoph, P., & Luborsky, L. (1998). Changes in CCRT pervasiveness during psychotherapy. In L. Luborsky & P. Crits-Christoph (Eds.), Understanding transference: The Core Conflictual Relationship Theme method (pp. 151-163). Washington, D.C.: American Psychological Association.
- Crittenden, P. M. (1994). Peering into the black box: An exploratory treatise on the development of self in young children. In D. Cicchetti & S. Toth (Eds.), Disorders and dysfunctions of the self (Vol. 5, pp. 79-148). Rochester, NY: University of Rochester Press.
- Crittenden, P. M. (1995). Attachment and psychopathology. In S. Goldberg & R. Muir & J. Kerr (Eds.), Attachment theory: Social, developmental, and clinical perspectives (pp. 367-406). Hillsdale, NJ: The Analytic Press.
- Cronbach, L. J., & Furby, L. (1970). How we should measure "change" - or should we? Psychological Bulletin, 74(1), 68-80.
- Crowell, J. A., & Treboux, D. (1995). A review of adult attachment measures: Implications for theory and research. Social Development, 4(3), 294-327.
- Curtis, J. T., Silberschatz, G., Sampson, H., & Weiss, J. (1994). The Plan Formulation Method. Psychotherapy Research, 4(3&4), 197-207.
- Dahl, H. (1991). The key to understanding change: Emotions as appetitive wishes and beliefs about their fulfillment. In J. Safran & L. Greenberg (Eds.), Emotion, Psychotherapy & Change (pp. 130-165). New York: Guilford.
- Dahl, H. (1998). The voyage of El Rubiyat to the discovery of FRAMES. In R. F. Bornstein & J. M. Masling (Eds.), Empirical Studies of the Therapeutic Hour (pp. 179-227). Washington, DC: American Psychological Association.
- Dahl, H., & Teller, V. (1994). The characteristics, identification, and applications of FRAMES. Psychotherapy Research, 4(3&4), 253-276.
- Dance, K. A., & Neufeld, R. W. J. (1988). Aptitude-treatment interaction research in the clinical setting: A review of attempts to dispel the "patient uniformity" myth. Psychological Bulletin, 104(2), 192-213.

- Davies-Osterkamp, S., Strauss, B. M., & Schmitz, N. (1996). Interpersonal problems as predictors of symptom related treatment outcome in longterm psychotherapy. Psychotherapy Research, 6(3), 164-176.
- del Carmen, R., & Huffman, L. (1996). Epilogue: Bridging the gap between research on attachment and psychopathology. Journal of Consulting and Clinical Psychology, 64(2), 291-294.
- Derogatis, L. R. (1977). SCL-90-R: Administration, Scoring, and Procedures Manual. Minneapolis, MN: National Computer Systems.
- Derogatis, L. R., & Cleary, P. A. (1977). Confirmation of the dimensional structure of the SCL-90: A study in construct validation. Journal of Clinical Psychology, 33(4), 981-989.
- DeWald, P. A. (1972). The clinical assessment of structural change. Journal of the American Psychoanalytic Association, 20, 302-325.
- Diamond, D., Clarkin, J., Levine, H., Levy, K., Foelsch, P., & Yeomans, F. (1999). Borderline conditions and attachment: A preliminary report. Psychoanalytic Inquiry, 19(5), 831-884.
- Doidge, N. (1997). Empirical evidence of the efficacy of psychoanalytic psychotherapies and psychoanalysis: An overview. Psychoanalytic Inquiry, Supplement, 102-150.
- Dozier, M. (1990). Attachment organization and treatment use for adults with serious psychopathological disorders. Development and Psychopathology, 2, 47-60.
- Dozier, M., Cue, K. L., & Barnett, L. (1994). Clinicians as caregivers: Role of attachment organization in treatment. Journal of Consulting and Clinical Psychology, 62(4), 793-800.
- Dozier, M., Lomax, L., Tyrell, C. L., & Lee, S. W. (2001). The challenge of treatment for clients with dismissing states of mind. Attachment and Human Development, 3(1), 62-76.
- Dozier, M., Stovall, K. C., & Albus, K. E. (1999). Attachment and psychopathology in adulthood. In J. Cassidy & P. R. Shaver (Eds.), Handbook of attachment: Theory, research, and clinical applications (pp. 497-519). New York: Guilford Press.
- Dreher, A. U. (2000). Foundations for Conceptual Research in Psychoanalysis (E. Ristl, Trans.). Madison, CT: International Universities Press.

- Dreher, M., Mengele, U., & Krause, R. (2001). Affective indicators of the psychotherapeutic process: An empirical case study. Psychotherapy Research, *11*(1), 99-117.
- Eagle, M. (1997). Attachment and psychoanalysis. British Journal of Medical Psychology, *70*, 217-229.
- Eames, V., & Roth, A. (2000). Patient attachment orientation and the early working alliance - A study of patient and therapist reports of alliance quality and ruptures. Psychotherapy Research, *10*(4), 421-434.
- Ekman, P., & Rosenberg, E. (Eds.). (1997). What the Face Reveals. Oxford: Oxford University Press.
- Elkin, I., Shea, T., Watkins, J. T., Imber, S. D., Sotsky, S. M., Collins, J. F., Glass, D. R., Pilkonis, P. A., Leber, W. R., Docherty, J. P., Fiester, S. J., & Parloff, M. B. (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program. Archives of General Psychiatry, *46*, 971-982.
- Elliott, R., & Anderson, C. (1994). Simplicity and complexity in psychotherapy research. In R. L. Russell (Ed.), Reassessing Psychotherapy Research (pp. 65-113). New York: The Guilford Press.
- Ellis, A. (1957). Outcome of employing three techniques of psychotherapy. Journal of Clinical Psychology, *13*, 344-350.
- Endicott, J., & Spitzer, R. L. (1978). A diagnostic interview: The schedule for affective disorders and schizophrenia. Archives of General Psychiatry, *35*, 837-844.
- Engel, G. L. (1968a). Closing comments. Journal of the American Psychoanalytic Association, *16*, 228-230.
- Engel, G. L. (1968b). Some obstacles to the development of research in psychoanalysis. Journal of the American Psychoanalytic Association, *16*, 195-203.
- Erickson, M. F., Korfmacher, J., & Egeland, B. R. (1992). Attachments past and present: Implications for therapeutic intervention with mother-infant dyads. Development and Psychopathology, *4*, 495-507.

- Erle, J. B. (1979). An approach to the study of analyzability and analyses: The course of forty consecutive cases selected for supervised analysis. Psychoanalytic Quarterly, 48, 198-228.
- Erle, J. B., & Goldberg, D. A. (1979). Problems in the assessment of analyzability. Psychoanalytic Quarterly, 48, 48-84.
- Erle, J. B., & Goldberg, D. A. (1984). Observations on assessment of analyzability by experienced analysts. Journal of the American Psychoanalytic Association, 32, 715-737.
- Escoll, P. J. (1987). Psychoanalysis of young adults: An overview. Psychoanalytic Inquiry, 7(1), 5-38.
- Eugster, S. L., & Wampold, B. E. (1996). Systematic effects of participant role on evaluation of the psychotherapy session. Journal of Consulting and Clinical Psychology, 64(5), 1020-1028.
- Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. Journal of Consulting Psychology, 16, 319-324.
- Eysenck, H. J., & Eysenck, S. B. G. (1975). Manual of the Eysenck Personality Questionnaire (junior & adult). London: Hodder and Stoughton.
- Fenichel, O. (1930). Statistischer Bericht über die therapeutische Tätigkeit, 1920-1930. In O. Fenichel (Ed.), Zehn Jahre Berliner Psychoanalytisches Institut (pp. 13-19). Vienna: International Psychoanalytisches Verlag.
- Fenichel, O. (1941). Problems of psychoanalytic technique (D. Brunswick, Trans.). New York: The Psychoanalytic Quarterly.
- 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. Journal of Psychotherapy Practice and Research, 10(4), 262-268.
- First, M. B., Gibbon, M., Spitzer, R. L., Williams, J. B. W., & Benjamin, L. S. (1997). User's guide for the Structured Clinical Interview for DSM-IV Axis II Personality Disorder (SCID-II). Washington, DC: American Psychiatric Press.

- First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1995). The Structured Clinical Interview for DSM-III-R personality disorders (SCID-II): Part I. Description. Journal of Personality Disorders, *9*(2), 83-91.
- First, M. B., Spitzer, R. L., Gibbon, M., Williams, J. B. W., Davies, M., Borus, J., Howes, M. J., Kane, J., Pope, H. G., Jr., & Rounsaville, B. (1995). The Structured Clinical Interview for DSM-III-R personality disorders (SCID-II): Part II. Multi-site test-retest reliability study. Journal of Personality Disorders, *9*(2), 92-104.
- Fisher, S., & Greenberg, R. P. (1977). The Scientific Credibility of Freud's Theories and Therapy. New York: Basic Books.
- Fisher, S., & Greenberg, R. P. (1996). Freud Scientifically Reappraised: Testing the Theories and Therapy. New York: John Wiley & Sons, Inc.
- Fiske, D. W., Hunt, H. F., Luborsky, L., Orne, M. T., Parloff, M. B., Reiser, M. F., & Tuma, A. H. (1970). Planning of research in effectiveness of psychotherapy. Archives of General Psychiatry, *22*, 22-32.
- Flemenbaum, A., & Zimmermann, R. L. (1973). Inter- and intra-rater reliability of the Brief Psychiatric Rating Scale. Psychological Reports, *36*, 783-792.
- Flick, S. N. (1988). Managing attrition in clinical research. Clinical Psychology Review, *8*, 499-515.
- Fonagy, P. (1998). Prevention, the appropriate target of infant psychotherapy. Infant Mental Health Journal, *19*(2), 124-150.
- Fonagy, P. (1999a). The justification of effectiveness studies in psychoanalysis. In IPA (Ed.), An open door review of outcome studies in psychoanalysis (pp. 56-64). London: Research Committee of the International Psychoanalytic Association.
- Fonagy, P. (1999b). The process of change and the change of processes: What can change in a 'good' analysis. Keynote address to the Spring Meeting of Division 39 of the American Psychological Association, New York April 16 1999.

- Fonagy, P. (1999c). Psychoanalytic theory from the viewpoint of attachment theory and research. In J. Cassidy & P. R. Shaver (Eds.), Handbook of attachment: Theory, research, and clinical applications (pp. 595-624). New York: Guilford Press.
- Fonagy, P. (1999d). Reflections on psychoanalytic research problems - an Anglo-Saxon view. In IPA (Ed.), An open door review of outcome studies in psychoanalysis (pp. 20-55). London: Research Committee of the International Psychoanalytic Association.
- Fonagy, P. (2001a). The justification of effectiveness studies in psychoanalysis, An open door review of outcome studies in psychoanalysis (Second ed., pp. 32-35). London: International Psychoanalytic Association.
- Fonagy, P. (2001b). Methodological considerations in evaluating the outcome of psychoanalysis, An open door review of outcome studies in psychoanalysis (Second ed., pp. 36-43). London: International Psychoanalytic Association.
- Fonagy, P. (2001c). Psychoanalytic assessment instruments, An open door review of outcome studies in psychoanalysis (Second ed., pp. 44-60). London: International Psychoanalytic Association.
- Fonagy, P. (2001d). Reflective Functioning (RF) Scale. In P. Fonagy (Ed.), An open door review of outcome studies in psychoanalysis (Second ed., pp. 60-62). London: International Psychoanalytic Association.
- Fonagy, P. (2001e). The talking cure in the cross fire of empiricism - the struggle for the hearts and minds of psychoanalytic clinicians: Commentary on papers by Lester Luborsky and Hans H. Strupp. Psychoanalytic Dialogues, 11(4), 647-658.
- Fonagy, P., Jones, E. E., Kächele, H., Krause, R., Clarkin, J., Perron, R., Gerber, A., & Allison, E. (2001). An open door review of outcome studies in psychoanalysis (Second ed.). London: International Psychoanalytic Association.
- Fonagy, P., Kächele, H., Krause, R., Jones, E. E., & Perron, R. (1999). An open door review of outcome studies in psychoanalysis: report prepared by the research committee of the IPA at the request of the President. London: International Psychoanalytic Association.

- Fonagy, P., Leigh, T., Steele, M., Steele, H., Kennedy, R., Mattoon, G., Target, M., & Gerber, A. (1996). The relation of attachment status, psychiatric classification, and response to psychotherapy. Journal of Consulting and Clinical Psychology, *64*, 22-31.
- Fonagy, P., Steele, H., & Steele, M. (1991). Maternal representations of attachment during pregnancy predict the organization of infant-mother attachment at one year of age. Child Development, *62*, 891-905.
- Fonagy, P., Steele, M., Steele, H., Leigh, T., Kennedy, R., Mattoon, G., & Target, M. (1995). Attachment, the reflective self, and borderline states: The predictive specificity of the Adult Attachment Interview and pathological emotional development. In S. Goldberg & R. Muir & J. Kerr (Eds.), Attachment theory: Social, developmental, and clinical perspectives. Hillsdale, NJ: The Analytic Press.
- Fonagy, P., Steele, M., Steele, H., Moran, G. S., & Higgitt, A. C. (1991). The capacity for understanding mental states: The reflective self in parent and child and its significance for security of attachment. Infant Mental Health Journal, *13*(3), 200-217.
- Fonagy, P., Steele, M., Steele, H., & Target, M. (1997). Reflective-Functioning Manual, version 4.1, for Application to Adult Attachment Interviews. London: University College London.
- Fonagy, P., & Tallandini-Shallice, M. (1993). On some problems of psychoanalytic research in practice. Bulletin of the Anna Freud Centre, *16*, 5-22.
- Fonagy, P., & Target, M. (1996). Predictors of outcome in child psychoanalysis: A retrospective study of 763 cases at the Anna Freud Centre. Journal of the American Psychoanalytic Association, *44*, 27-77.
- Fonagy, P., & Target, M. (1997). Attachment and reflective function: Their role in self-organization. Development and Psychopathology, *9*, 679-700.
- Fonagy, P., Target, M., Cottrell, D., Phillips, J., & Kurtz, Z. (2002). What Works for Whom? A critical review of treatments for children and adolescents. New York: Guilford.
- Fox, N. (1995). Of the way we were: Adult memories about attachment experiences and their role in determining infant-parent relationships: A commentary on van IJzendoorn (1995). Psychological Bulletin, *117*(3), 404-410.

- Frances, A. (1982). Personality change within and outside psychoanalysis. Journal of the American Academy of Psychoanalysis, *10*, 527-539.
- Frances, A., Clarkin, J. F., Gilmore, M., Hurt, S. W., & Brown, R. (1984). Reliability of criteria for borderline personality disorder: A comparison of DSM-III and the Diagnostic Interview for Borderline patients. American Journal of Psychiatry, *141*(9), 1080-1084.
- Froyd, J. E., Lambert, M. J., & Froyd, J. D. (1996). A review of practices of psychotherapy outcome measurement. Journal of Mental Health, *5*(1), 11-15.
- Gabbard, G. O., Coyne, L., Kennedy, L. L., Beasley, C., Deering, C. D., Schroder, P., Larson, J., & Cerney, M. S. (1987). Interrater reliability in the use of the Brief Psychiatric Rating Scale. Bulletin of the Menninger Clinic, *51*(6), 519-531.
- Gabbard, G. O., Gunderson, J. G., & Fonagy, P. (2002). The place of psychoanalytic treatments within psychiatry. Archives of General Psychiatry, *59*(6), 505-510.
- Gabbard, G. O., Horwitz, L., Allen, J. G., Frieswyk, S., Newsom, G., Colson, D. B., & Coyne, L. (1994). Transference interpretation in the psychotherapy of borderline patients: A high-risk, high-gain phenomenon. Harvard Review of Psychiatry, *2*, 59-69.
- Gabbard, G. O., Lazar, S. G., Hornberger, J., & Spiegel, D. (1997). The economic impact of psychotherapy: A review. American Journal of Psychiatry, *154*(2), 147-155.
- Galatzer-Levy, R. M., Bachrach, H., Skolnikoff, A., & Waldron, S., Jr. (2000). Does Psychoanalysis Work? New Haven: Yale University Press.
- Garduk, E. L., & Haggard, E. A. (1972). Immediate effects on patients of psychoanalytic interpretations. Psychological Issues, *7*(4).
- Garfield, S. L. (1982). Eclecticism and integration in psychotherapy. Behavior Therapy, *13*, 610-623.
- Garfield, S. L. (1994). Research on client variables in psychotherapy. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of psychotherapy and behavior change (Fourth ed., pp. 190-228). New York: John Wiley & Sons, Inc.
- Garfield, S. L. (1998). Some comments on empirically supported treatments. Journal of Consulting and Clinical Psychology, *66*, 121-125.

- Garfield, S. L., & Bergin, A. E. (1994). Introduction and historical overview. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of psychotherapy and behavior change (Fourth ed., pp. 3-18). New York: John Wiley & Sons, Inc.
- Gaston, L. (1990). The concept of the alliance and its role in psychotherapy: Theoretical considerations and empirical considerations. Psychotherapy, *27*(2), 143-153.
- Gaston, L. (1991). Reliability and criterion-related validity of the California Psychotherapy Alliance Scales-Patient Version. Psychological Assessment, *3*(1), 68-74.
- Gaston, L., Marmar, C. R., Thompson, L. W., & Gallagher, D. (1988). Relation of patient pretreatment characteristics to the therapeutic alliance in diverse psychotherapies. Journal of Consulting and Clinical Psychology, *56*(4), 483-489.
- Gaston, L., Piper, W. E., Debbane, E. G., Bienvenu, J. P., & Garant, J. (1994). Alliance and technique for predicting outcome in short- and long-term analytic psychotherapy. Psychotherapy Research, *4*, 121-135.
- Gaston, L., & Ring, J. M. (1992). Preliminary results on the Inventory of Therapeutic Strategies. Journal of Psychotherapy Practice and Research, *4*(2), 135-146.
- Gaston, L., Thompson, L., Gallagher, D., Cournoyer, L.-G., & Gagnon, R. (1998). Alliance, technique, and their interactions in predicting outcome of behavioral, cognitive, and brief dynamic therapy. Psychotherapy Research, *8*(2), 190-209.
- George, C., Kaplan, N., & Main, M. (1996). Adult Attachment Interview. Unpublished manuscript. (Third ed.). Berkeley, CA: University of California.
- George, C., & Solomon, J. (1999). Attachment and caregiving: The caregiving behavioral system. In J. Cassidy & P. R. Shaver (Eds.), Handbook of attachment: Theory, research, and clinical applications (pp. 649-670). New York: Guilford Press.
- Gerber, A. J. (1994). Problems in assessing adolescent psychopathology and adaptation: An assessment of cross-informant correspondence of the Child Behavior Checklist in an outpatient adolescent sample. Unpublished M.Sc. thesis, University College London, London, England.

- Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., Paus, T., Evans, A. C., & Rapoport, J. L. (1999). Brain development during childhood and adolescence: a longitudinal MRI study. Nature Neuroscience, *2*(10), 861-863.
- Gill, M. M. (1979). The analysis of the transference. Journal of the American Psychoanalytic Association, *24*, 263-288.
- Goldfried, M. R., Greenberg, L. S., & Marmar, C. (1990). Individual psychotherapy: Process and outcome. Annual Review of Psychology, *41*, 659-688.
- Goldfried, M. R., & Wolfe, B. E. (1998). Towards a more clinically valid approach to therapy research. Journal of Consulting and Clinical Psychology, *66*, 143-150.
- Gomes-Schwartz, B. (1978). Effective ingredients in psychotherapy: Prediction of outcome from process variables. Journal of Consulting and Clinical Psychology, *46*(5), 1023-1035.
- Gomes-Schwartz, B., & Schwartz, J. M. (1978). Psychotherapy process variables distinguishing the "inherently helpful" person from the professional psychotherapist. Journal of Consulting and Clinical Psychology, *46*(1), 196-197.
- Graff, H., & Luborsky, L. (1977). Long-term trends in transference and resistance: A report on a quantitative-analytic method applied to four psychoanalyses. Journal of the American Psychoanalytic Association, *25*, 471-486.
- Grande, T., Rudolf, G., Oberbracht, C., Jakobsen, T., & Keller, W. (2004). Investigating structural change in the process and outcome of a psychoanalytic treatment: The Heidelberg-Berlin Study. In P. Richardson & H. Kächele & C. Renlund (Eds.), Research on Psychoanalytic Psychotherapy with Adults. London: Karnac.
- Grant, J., & Sandell, R. (2004). Close family or mere neighbours? Some empirical data on the differences between psychoanalysis and psychotherapy. In P. Richardson & H. Kächele & C. Renlund (Eds.), Research on Psychoanalytic Psychotherapy with Adults (pp. 81-108). London: Karnac.
- Greenbaum, P. E., Dedrick, R. F., Prange, M. E., & Friedman, R. M. (1994). Parent, teacher and child ratings of problem behaviors of youngsters with serious emotional disturbances. Psychological Assessment, *6*, 141-148.

- Greenberg, L. S. (1986). Change process research. Journal of Consulting and Clinical Psychology, 54(1), 4-9.
- Greenberg, L. S. (1991). Research on the process of change. Psychotherapy Research, 1(1), 3-16.
- Greenson, R. R. (1994). The technique and practice of psychoanalysis. London: Hogarth Press.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Moran (Eds.), Syntax and semantics III: Speech acts (pp. 41-58). New York: Academic Press.
- Grossmann, K. E. (1995). The evolution and history of attachment research and theory. In S. Goldberg & R. Muir & J. Kerr (Eds.), Attachment theory: Social, developmental, and clinical perspectives (pp. 85-121). Hillsdale, NJ: Analytic Press.
- Gunderson, J. G., Kolb, J. E., & Austin, V. (1981). The Diagnostic Interview for Borderline patients. American Journal of Psychiatry, 137(7), 896-903.
- Gunderson, J. G., Najavits, L. M., Leonhard, C., Sullivan, C. N., & Sabo, A. N. (1997). Ontogeny of the therapeutic alliance in borderline patients. Psychotherapy Research, 7(3), 301-309.
- Hageman, W. J. J. M., & Arrindell, W. A. (1999a). Clinically significant and practical! Enhancing precision does make a difference. Reply to McGlinchey and Jacobso, Hsu, and Speer. Behaviour Research and Therapy, 37, 1219-1233.
- Hageman, W. J. J. M., & Arrindell, W. A. (1999b). Establishing clinically significant change: increment of precision and the distinction between individual and group level of analysis. Behaviour Research and Therapy, 37, 1169-1193.
- Hamburg, D. A., Bibring, G. L., Fisher, C., Stanton, A., Wallerstein, R. S., Weinstock, H. I., & Haggard, E. (1967). Report of Ad Hoc Committee on Central Fact-Gathering Data of the American Psychoanalytic Association. Journal of the American Psychoanalytic Association, 15, 841-861.
- Hardy, G., Aldridge, J., Davidson, C., Rowe, C., Reilly, S., & Shapiro, D. A. (1999). Therapist responsiveness to client attachment styles and issues observed in client-identified significant events in psychodynamic-interpersonal psychotherapy. Psychotherapy Research, 9(1), 36-53.

- Harris, T. O. (1997). Adult attachment processes and psychotherapy: A commentary on Bartholomew and Birtchnell. British Journal of Medical Psychology, 70, 281-290.
- Harrison, A., M. (1998). The case of Sophie. Infant Mental Health Journal, 19(3), 309-314.
- Hartley, D., & Strupp, H. H. (1983). The therapeutic alliance: Its relationship to outcome in brief psychotherapy. In J. M. Masling (Ed.), Empirical Studies of Psychoanalytic Theories (Vol. 1, pp. 1-27). Hillsdale, NJ: Erlbaum.
- Hasin, D. S., & Grant, B. F. (1987a). Diagnosing depressive disorders in patients with alcohol and drug problems: A comparison of the SADS-L and the DIS. Journal of Psychiatric Research, 21(3), 301-311.
- Hasin, D. S., & Grant, B. F. (1987b). Psychiatric diagnosis of patients with substance abuse problems: A comparison of two procedures, the DIS and the SADS-L. Journal of Psychiatric Research, 21(1), 7-22.
- Hatcher, R. L., & Barends, A. (1996). Patients' view of the alliance in psychotherapy: exploratory factor analysis of three alliance measures. Journal of Consulting and Clinical Psychology, 64, 1326-1336.
- Heilbrunn, G. (1966). Results with psychoanalytic therapy and professional commitment. American Journal of Psychotherapy, 20, 89-99.
- Heine, R. W. (1953). A comparison of patients' reports on psychotherapeutic experience with psychoanalytic, nondirective, and Adlerian therapists. American Journal of Psychotherapy, 7, 16-23.
- Heinicke, C. M. (1965). Frequency of psychotherapy session as a factor affecting the child's developmental status. Psychoanalytic Study of the Child, 20, 2-98.
- Heinicke, C. M., & Ramsey-Klee, D. M. (1986). Outcome of child psychotherapy as a function of frequency of session. Journal of the American Academy of Child Psychiatry, 25(2), 247-253.
- Henry, W. P., Schacht, T. E., & Strupp, H. H. (1986). Structural analysis of social behavior: Application to a study of interpersonal process in differential psychotherapeutic outcome. Journal of Consulting and Clinical Psychology, 54(1), 27-31.

- Henry, W. P., Strupp, H. H., Schacht, T. E., & Gaston, L. (1994). Psychodynamic approaches. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of psychotherapy and behavior change (Fourth ed., pp. 467-508). New York: John Wiley & Sons, Inc.
- Hesse, E. (1996). Discourse, memory and the Adult Attachment Interview: A note with emphasis on the emerging Cannot Classify category. Infant Mental Health Journal, 17, 4-11.
- Hesse, E. (1999). The Adult Attachment Interview: Historical and current perspectives. In J. Cassidy & P. R. Shaver (Eds.), Handbook of attachment: Theory, research, and clinical applications (pp. 395-433). New York: Guilford Press.
- Hill, C. E. (1986). An overview of the Hill Counselor and Client Verbal Response Mode Category Systems. In L. S. Greenberg & W. M. Pinsof (Eds.), The psychotherapeutic process: A research handbook (pp. 131-159). New York: Guilford.
- Hill, C. E., Helms, J. E., Tichenor, V., Spiegel, S. B., O'Grady, K. E., & Perry, E. S. (1988). Effects of therapist response modes in brief psychotherapy. Journal of Counseling Psychology, 35(3), 222-233.
- Hill, C. E., O'Grady, K. E., & Elkin, I. (1992). Applying the Collaborative Study Psychotherapy Rating Scale to rate adherence in cognitive-behavior therapy, interpersonal therapy, and clinical management. Journal of Consulting and Clinical Psychology, 60, 73-79.
- Hilliard, R. B. (1993). Single-case methodology in psychotherapy process and outcome research. Journal of Consulting and Clinical Psychology, 61(3), 373-380.
- Høglend, P. (1993a). Personality disorders and long-term outcome after brief dynamic psychotherapy. Journal of Personality Disorders, 7(2), 168-181.
- Høglend, P. (1993b). Suitability for brief dynamic psychotherapy: Psychodynamic variables as predictors of outcome. Acta Psychiatrica Scandinavica, 88, 104-110.
- Høglend, P. (1993c). Transference interpretations and long-term change after dynamic psychotherapy of brief to moderate length. Special section: Transference interpretation. American Journal of Psychotherapy, 47, 494-507.

- Høglend, P. (1996). Long-term effects of transference interpretations: comparing results from a quasi-experimental and a naturalistic long-term follow-up study of brief dynamic psychotherapy. Acta Psychiatrica Scandinavica, *93*, 205-211.
- Høglend, P., Bøgwald, K.-P., Amlo, S., Heyerdahl, O., Sørbye, Ø., Marble, A., Sjaastad, M. C., & Bentsen, H. (2000). Assessment of change in dynamic psychotherapy. Journal of Psychotherapy Practice and Research, *9*(4), 190-199.
- Høglend, P., Heyerdahl, O., Amlo, S., Engelstad, V., Fossum, A., Sørbye, Ø., & Sørli, T. (1993). Interpretations of the patient-therapist relationship in brief dynamic psychotherapy. Journal of Psychotherapy Practice and Research, *2*, 296-306.
- Høglend, P., & Piper, W. E. (1995). Focal adherence in brief dynamic psychotherapy: A comparison of findings from two independent studies. Psychotherapy, *32*(4), 618-628.
- Høglend, P., Sørli, T., Heyerdahl, O., Sørbye, Ø., & Amlo, S. (1993). Brief dynamic psychotherapy: Patient suitability, treatment length, and outcome. Journal of Psychotherapy Practice and Research, *2*, 230-241.
- Høglend, P., Sørli, T., Sørbye, Ø., Heyerdahl, O., & Amlo, S. (1992). Long-term changes after brief dynamic psychotherapy: symptomatic versus dynamic assessments. Acta Psychiatrica Scandinavica, *86*, 165-172.
- Holland, S. J., Roberts, N. E., & Messer, S. B. (1998). Reliability and validity of the Rutgers Psychotherapy Process Scale. Psychotherapy Research, *8*(1), 104-110.
- Hollon, S. D., & Flick, S. N. (1988). On the meaning and methods of clinical significance. Behavioral Assessment, *10*, 197-206.
- Holmes, J. (1997). Attachment, autonomy, intimacy: Some clinical implications of attachment theory. British Journal of Medical Psychology, *70*, 231-248.
- Holmes, J. (1998). The changing aims of psychoanalytic psychotherapy: An integrative perspective. International Journal of Psycho-analysis, *79*(2), 227-240.
- Holtzworth-Munroe, A., Stuart, G. L., & Hutchinson, G. (1997). Violent versus nonviolent husbands: Differences in attachment patterns, dependency, and jealousy. Journal of Family Psychology, *11*(3), 314-331.

- Hölzer, M., Pokorny, D., Kächele, H., & Luborsky, L. (1997). The verbalization of emotion in the therapeutic dialogue - A correlate of therapeutic outcome? Psychotherapy Research, *7*(3), 261-273.
- Horowitz, L. M., & Rosenberg, S. E. (1994). The Consensual Response Psychodynamic Formulation: Part 1: Method and research results. Psychotherapy Research, *4*(3&4), 222-233.
- Horowitz, L. M., Rosenberg, S. E., Baer, B. A., Ureño, G., & Villaseñor, V. S. (1988). Inventory of interpersonal problems: Psychometric properties and clinical applications. Journal of Consulting and Clinical Psychology, *56*(6), 885-892.
- Horowitz, L. M., Rosenberg, S. E., & Bartholomew, K. (1993). Interpersonal problems, attachment styles, and outcome in brief dynamic psychotherapy. Journal of Consulting and Clinical Psychology, *61*(4), 549-560.
- Horowitz, L. M., Rosenberg, S. E., Ureño, G., Kalehzan, B. M., & O'Halloran, P. (1989). Psychodynamic formulation, consensual response method, and interpersonal problems. Journal of Consulting and Clinical Psychology, *57*(5), 599-606.
- Horowitz, L. M., Strupp, H. H., Lambert, M. J., & Elkin, I. (1997). Overview and summary of the core battery conference. In H. H. Strupp & L. M. Horowitz & M. J. Lambert (Eds.), Measuring Patient Changes in Mood, Anxiety, and Personality Disorders: Toward a core battery (pp. 11-54). Washington, DC: American Psychological Association.
- Horowitz, M. (1993). Defensive control of states and person schemas. Journal of the American Psychoanalytic Association, *41* supplement, 67-89.
- Horowitz, M. J. (1989). Relationship schema formulation: Role-relationship models and intrapsychic conflict. Psychiatry, *52*, 260-274.
- Horowitz, M. J. (1994). Configurational analysis and the use of role-relationship models to understand transference. Psychotherapy Research, *4*(3&4), 184-196.
- Horowitz, M. J., Eells, T., Singer, J., & Salovey, P. (1995). Role-relationship models for case formulation. Archives of General Psychiatry, *52*, 625-632.

- Horowitz, M. J., & Eells, T. D. (1993). Case formulations using Role-Relationship Model Configurations: A reliability study. Psychotherapy Research, 3(1), 57-68.
- Horowitz, M. J., Ewert, M., & Milbrath, C. (1996). States of emotional control during psychotherapy. Journal of Psychotherapy Practice and Research, 5, 20-25.
- Horowitz, M. J., Kernberg, O. F., & Weinshel, E. M. (Eds.). (1993). Psychic Structure and Psychic Change: Essays in honor of Robert S. Wallerstein, M.D. Madison, CT: International Universities Press, Inc.
- Horowitz, M. J., Marmar, C. R., Weiss, D. S., Kaltreider, N. B., & Wilner, N. R. (1986). Comprehensive analysis of change after brief dynamic psychotherapy. American Journal of Psychiatry, 143(5), 582-589.
- Horowitz, M. J., & Stinson, C. H. (1995). Consciousness and the processes of control. Journal of Psychotherapy Practice and Research, 4(2), 123-139.
- Horvath, A. O. (1988). Placebos and common factors in two decades of psychotherapy research. Psychological Bulletin, 104, 214-225.
- Horvath, A. O., & Greenberg, L. (1986). The development of the Working Alliance Inventory. In L. S. Greenberg & W. M. Pinsof (Eds.), The psychotherapeutic process: A research handbook (pp. 529-556). New York: Guilford.
- Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. Journal of Counseling Psychology, 36(2), 223-233.
- Horvath, A. O., & Symonds, B. D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. Journal of Counseling Psychology, 38(2), 139-149.
- Horwitz, L. (1974). Clinical prediction in psychotherapy. New York: Jason Aronson.
- Horwitz, L., Hubbard, G. O., Allen, J. G., Frieswyk, S. H., Colson, D. B., Newsom, G. E., & Coyne, L. (1996). Borderline Personality Disorder: Tailoring the Psychotherapy to the Patient. Washington, DC: American Psychiatric Press, Inc.
- Howard, K. I., Kopta, S. M., Krause, M. S., & Orlinsky, D. E. (1986). The dose-effect relationship in psychotherapy. American Psychologist, 41, 159-164.

- Hsu, L. M. (1995). Regression toward the mean associated with measurement error and the identification of improvement and deterioration in psychotherapy. Journal of Consulting and Clinical Psychology, *63*(1), 141-144.
- Hsu, L. M. (1999). A comparison of three methods of identifying reliable and clinically significant changes: commentary on Hageman and Arrindell. Behaviour Research and Therapy, *37*, 1195-1202.
- Huber, D., & Klug, G. (2004). Contribution to the measurement of mode-specific effects in long-term psychoanalytic psychotherapy. In P. Richardson & H. Kächele & C. Renlund (Eds.), Research on Psychoanalytic Psychotherapy with Adults (pp. 63-80). London: Karnac.
- Imber, S. D. (1992). Then and now: Forty years in psychotherapy research. Clinical Psychology Review, *12*, 199-204.
- Ingram, D. H. (1983). Clinical indicators of structural change. American Journal of Psychoanalysis, *43*(1), 39-48.
- IPA. (2001a). Core conflictual relationship theme method (CCRT). In P. Fonagy (Ed.), An open door review of outcome studies in psychoanalysis (Second ed., pp. 182-183). London: International Psychoanalytic Association.
- IPA. (2001b). The Saarbrücken studies on unconscious interaction regulation: Multi-channel psychotherapy process research projects (SSUIR). In P. Fonagy (Ed.), An open door review of outcome studies in psychoanalysis (Second ed., pp. 228-232). London: International Psychoanalytic Association.
- IPA. (2001c). The Stuttgart study: The Stuttgart psychotherapeutic hospital follow-up study (TSS). In P. Fonagy (Ed.), An open door review of outcome studies in psychoanalysis (Second ed., pp. 106-109). London: International Psychoanalytic Association.
- IPA. (2001d). Ulm-Leipzig-Göttingen studies of transference patterns using the CCRT method. In P. Fonagy (Ed.), An open door review of outcome studies in psychoanalysis (Second ed., pp. 184-185). London: International Psychoanalytic Association.
- Jacobs, T. J., & Chused, J. F. (1987). Psychoanalysis of the young adult: Theory and technique. Journal of the American Psychoanalytic Association, *35*, 175-186.

- Jacobson, N. S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. Behavior Therapy, 15, 336-352.
- Jacobson, N. S., Roberts, L. J., Berns, S. B., & McGlinchey, J. B. (1999). Methods for defining and determining the clinical significance of treatment effects: Description, application, and alternatives. Journal of Consulting and Clinical Psychology, 67(3), 300-307.
- 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(1), 12-19.
- Jadad, A. (1998). Randomised controlled trials. London: BMJ Books.
- Johnson, M. E., Popp, C., Schacht, T. E., Mellon, J., & Strupp, H. H. (1989). Converging evidence for identification of recurrent relationship themes: Comparison of two methods. Psychiatry, 52, 275-287.
- Jones, E. (1936). Decannual Report of the London Clinic of Psychoanalysis. 1926-36. London: London Clinic of Psychoanalysis.
- Jones, E. E. (1993a). How will psychoanalysis study itself? Journal of the American Psychoanalytic Association, 41 supplement, 91-108.
- Jones, E. E. (1993b). Introduction to special section: Single-case research in psychotherapy. Journal of Consulting and Clinical Psychology, 61(3), 371-372.
- Jones, E. E. (1997). Modes of therapeutic action. International Journal of Psycho-Analysis, 78(6), 1135-1150.
- Jones, E. E. (2000). Therapeutic Action: A Guide to Psychoanalytic Therapy. Northvale, NJ: Jason Aronson.
- Jones, E. E., Cumming, J. D., & Pulos, S. M. (1993). Tracing clinical themes across phases of treatment by a Q-set. In N. E. Miller & L. Luborsky & J. P. Barber & J. P. Docherty (Eds.), Psychodynamic Treatment Research: A Handbook for Clinical Practice (pp. 14-36). New York: Basic Books.

- Jones, E. E., Ghannam, J., Nigg, J. T., & Dyer, J. F. P. (1993). A paradigm for single-case research: The time series study of a long-term psychotherapy for depression. Journal of Consulting and Clinical Psychology, *61*(3), 381-394.
- Jones, E. E., Hall, S. A., & Parke, L. A. (1991). The process of change: The Berkeley Psychotherapy Research Group. In L. E. Beutler & M. Crago (Eds.), Psychotherapy Research: An International Review of Programmatic Studies (pp. 98-106). Washington, D. C.: American Psychological Association.
- Jones, E. E., & Price, P. B. (1998). Interaction structure and change in psychoanalytic therapy. In R. F. Bornstein & J. M. Masling (Eds.), Empirical Studies of the Therapeutic Hour (pp. 27-62). Washington, DC: American Psychological Association.
- Jones, E. E., & Pulos, S. M. (1993). Comparing the process in psychodynamic and cognitive-behavioral therapies. Journal of Consulting and Clinical Psychology, *61*(2), 306-316.
- Jones, E. E., & Windholz, M. (1990). The psychoanalytic case study: Toward a method for systematic inquiry. Journal of the American Psychoanalytic Association, *38*, 985-1012.
- Juffer, F., van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (1997). Intervention in transmission of insecure attachment: A case study. Psychological Reports, *80*, 531-543.
- Kächele, H. (1992). Narration and observation in psychotherapy research: Reporting on a 20 year long journey from qualitative case reports to quantitative studies on the psychoanalytic process. Psychotherapy Research, *2*(1), 1-15.
- Kaltreider, N. B., DeWitt, K. N., Weiss, D. S., & Horowitz, M. J. (1981). Patterns of individual change scales. Archives of General Psychiatry, *38*, 1263-1269.
- Kantrowitz, J. L. (1993). Outcome research in psychoanalysis: Review and reconsiderations. Journal of the American Psychoanalytic Association, *41* supplement, 313-328.
- Kantrowitz, J. L. (1997). A brief review of psychoanalytic outcome research. Psychoanalytic Inquiry, Supplement, 87-101.
- Kantrowitz, J. L., Katz, A. L., Paolitto, F., Sashin, J., & Solomon, L. (1987). Changes in the level and quality of object relations in psychoanalysis: Followup of a longitudinal, prospective study. Journal of the American Psychoanalytic Association, *35*, 23-46.

- Kanzer, M. (1968). Discussion of Engel 1968. Journal of the American Psychoanalytic Association, 16, 211-215.
- Kaplan, H. I., Sadock, B. J., & Grebb, J. A. (1994). Synopsis of psychiatry. Baltimore: William & Wilkins.
- Karush, A. (1956). People who seek help: a diagnostic survey. In S. Rado & G. E. Daniels (Eds.), Changing Concepts of Psychoanalytic Medicine (pp. 69-84). New York: Grune & Stratton.
- Kashani, J. H., Orvaschel, H., Burk, J. P., & Reid, J. C. (1985). Informant variance: The issue of parent-child disagreement. Journal of the American Academy of Child and Adolescent Psychiatry, 24, 437-441.
- Kazdin, A. E. (1982). Single-case experimental designs in clinical research and practice. New Directions for Methodology of Social and Behavioral Science, 13, 33-47.
- Kazdin, A. E. (1986). Comparative outcome studies of psychotherapy: Methodological issues and strategies. Journal of Consulting and Clinical Psychology, 54(1), 95-105.
- Kazdin, A. E. (Ed.). (1992). Methodological issues and strategies in clinical research. Washington DC: American Psychological Association.
- Kazdin, A. E. (1994). Methodology, design, and evaluation in psychotherapy research. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of psychotherapy and behavior change (Fourth ed., pp. 19-71). New York: John Wiley & Sons, Inc.
- Kazdin, A. E. (1999). The meanings and measurement of clinical significance. Journal of Consulting and Clinical Psychology, 67(3), 332-339.
- Kendall, P. C. (1998). Empirically supported psychological therapies: Introduction to special section. Journal of Consulting and Clinical Psychology, 66(1), 3-6.
- Kendall, P. C., & Grove, W. M. (1988). Normative comparisons in therapy outcome. Behavioral Assessment, 10, 147-158.
- Kendall, P. C., Marrs-Garcia, A., Nath, S. R., & Sheldrick, R. C. (1999). Normative comparisons for the evaluation of clinical significance. Journal of Consulting and Clinical Psychology, 67(3), 285-299.

- Kennedy, H., & Moran, G. (1991). Reflections on the aim of child analysis. Psychoanalytic Study of the Child, 46, 181-198.
- Kernberg, O. (1999). Psychoanalysis, psychoanalytic psychotherapy and supportive psychotherapy: Contemporary controversies. International Journal of Psychoanalysis, 80, 1075-1091.
- Kernberg, O. F. (1973). Summary and conclusions of "Psychotherapy and psychoanalysis, final report of the Menninger Foundation's Psychotherapy Research Project". International Journal of Psychiatry, 11, 62-77.
- Kernberg, O. F. (1976). Object relations theory and clinical psychoanalysis. New York: Aronson.
- Kernberg, O. F., Burstein, E. D., Coyne, L., Appelbaum, A., Horwitz, L., & Voth, H. (1972). Psychotherapy and psychoanalysis: Final report of the Menninger Foundation's Psychotherapy Research Project. Bulletin of the Menninger Clinic, 36(1/2), 1-275.
- Kessel, L., & Hyman, H. (1933). The value of psychoanalysis as a therapeutic procedure. Journal of the American Medical Association, 101, 1612-1615.
- Kiesler, D. J. (1966). Some myths of psychotherapy research and the search for a paradigm. Psychological Bulletin, 65(2), 110-136.
- Kiesler, D. J. (1995). Research classic: "Some myths of psychotherapy research and the search for a paradigm": revisited. Psychotherapy Research, 5(2), 91-101.
- Klein, M. (1937/1984). Love, guilt, and reparation, Love, Guilt, and Reparation: The Writings of Melanie Klein Volume I (pp. 306-343). New York: Macmillan.
- Klein, M. (1948). Contributions to psychoanalysis, 1921-1945. London: Hogarth Press.
- Knapp, P. H., Levin, S., McCarter, R. H., Wermer, H., & Zetzel, E. (1960). Suitability for psychoanalysis: A review of one hundred supervised analytic cases. Psychoanalytic Quarterly, 29, 459-477.
- Knight, R. (1941). Evaluation of the results of psychoanalytic therapy. American Journal of Psychiatry, 98, 434-436.
- Kohut, H. (1977). The Restoration of the Self. New York: International Universities Press.

- Kolden, G. G. (1991). The generic model of psychotherapy: an empirical investigation of patterns of process and outcome relationships. Psychotherapy Research, *1*, 62-73.
- Kolden, G. G. (1996a). Change in early sessions of dynamic therapy: Universal processes and the generic model of psychotherapy. Journal of Consulting and Clinical Psychology, *64*(3), 489-496.
- Kolden, G. G. (1996b). Effective microprocesses in early sessions of dynamic psychotherapy. Journal of psychotherapy practice and research, *5*, 122-131.
- Kolden, G. G., & Howard, K. I. (1992). An empirical test of the generic model of psychotherapy. Journal of Psychotherapy Practice and Research, *1*(3), 225-236.
- Kolden, G. G., & Klein, M. (1996). Therapeutic process in dynamic therapy for personality disorders: The joint influence of acute distress and dysfunction and severity of personality pathology. Journal of Personality Disorders, *10*(2), 107-121.
- Kordy, H., von Rad, M., & Senf, W. (1983). Success and failure in psychotherapy: Hypotheses and results from the Heidelberg Follow-Up Project. Psychotherapy and Psychosomatics, *40*, 211-227.
- Kordy, H., von Rad, M., & Senf, W. (1988). Time and its relevance for a successful psychotherapy. Psychotherapy and Psychosomatics, *49*, 212-222.
- Kordy, H., von Rad, M., & Senf, W. (1989). Empirical hypotheses on the psychotherapeutic treatment of psychosomatic patients in short and long-term time-unlimited psychotherapy. Psychotherapy and Psychosomatics, *52*, 155-163.
- Korfmacher, J., Adam, E. K., Ogawa, J., & Egeland, B. (1997). Adult attachment: Implications for the therapeutic process in a home visitation intervention. Applied Developmental Science, *1*(1), 43-52.
- Krause, R., Steimer-Krause, E., Merten, J., & Ullrich, B. (1998). Dyadic interaction regulation emotion and psychopathology. In W. Flack & J. Laird (Eds.), Emotions and Psychopathology: Theory and Research (pp. 70-80). Oxford: Oxford University Press.
- Lambert, M. J. (1999). Are differential treatment effects inflated by researcher therapy allegiance? Could Clever Hans count? Clinical Psychology: Science and Practice, *6*(1), 127-130.

- Lambert, M. J., & Hill, C. E. (1994). Assessing psychotherapy outcomes and processes. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of psychotherapy and behavior change (Fourth ed., pp. 72-113). New York: John Wiley & Sons, Inc.
- Lazar, S. G. (Ed.). (1997). Extended Dynamic Psychotherapy: Making the case in an era of managed care. Hillsdale, New Jersey: Analytic Press.
- Leary, K., & Gutfreund, M. J. (1992). The vicissitudes of transference interpretations: The transference talk scale. Paper presented at the Annual meeting of the International Society for Psychotherapy Research, Lyon, France.
- Leuzinger-Bohleber, M. (1999, July 24, 1999). Longterm effects of psychoanalyses and psychoanalytic therapies: A representative follow-up study. Paper presented at the Empirical Precongress of the IPA Congress, Santiago, Chile.
- Leuzinger-Bohleber, M., & Target, M. (Eds.). (2002). Outcomes of Psychoanalytic Treatment: Perspectives for Therapists and Researchers. New York: Whurr.
- Lichtenberg, J. (2003). A clinician's view of attachment theory and research: A discussion of the papers in three issues of Psychoanalytic Inquiry. Psychoanalytic Inquiry, 23(1), 151-206.
- Liotti, G. (1995). Disorganized/disoriented attachment in the psychotherapy of dissociative disorders. In S. Goldberg & R. Muir & J. Kerr (Eds.), Attachment theory: Social, developmental, and clinical perspectives (pp. 343-363). Hillsdale, NJ: The Analytic Press.
- Little, R. J., & Rubin, D. B. (1987). Statistical analysis with missing data. New York, NY: Wiley.
- Luborsky, L. (1984). Principles of Psychoanalytic Psychotherapy: A Manual for Supportive-Expressive Treatment. New York: Basic Books, Inc.
- Luborsky, L. (1998). Alternative measures of the central relationship pattern. In L. Luborsky & P. Crits-Christoph (Eds.), Understanding transference: The Core Conflictual Relationship Theme method (Second ed., pp. 289-303). Washington, D.C.: American Psychological Association.
- Luborsky, L. (2000). An introduction to central relationship pattern measures: The Central Relationship Questionnaire. Journal of Psychotherapy Practice and Research, 9(4), 200.

- Luborsky, L. (2001). The meaning of Empirically Supported Treatment research for psychoanalytic and other long-term therapies. Psychoanalytic Dialogues, 11(4), 583-604.
- Luborsky, L., Bachrach, H., Graff, H., Pulver, S., & Christoph, P. (1979). Precondition and consequences of transference interpretations: A clinical quantitative investigation. Journal of Nervous and Mental Disease, 167, 391-401.
- Luborsky, L., & Barber, J. P. (1994). Perspectives on seven transference-related measures applied to the interview with Ms. Smithfield. Psychotherapy Research, 4(3&4), 152-154.
- Luborsky, L., Barber, J. P., Siqueland, L., Johnson, S., Najavits, L. M., Frank, A., & Daley, D. (1996). The Revised Helping Alliance Questionnaire (HAq-II). Journal of Psychotherapy Practice and Research, 5(3), 260-271.
- Luborsky, L., & Crits-Christoph, P. (1998). Understanding Transference: The Core Conflictual Relationship Theme Method (Second ed.). Washington DC: American Psychological Association.
- Luborsky, L., Crits-Christoph, 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(8), 480-491.
- Luborsky, L., Crits-Christoph, P., & Mellon, J. (1986). Advent of objective measures of the transference concept. Journal of Clinical and Consulting Psychology, 54(1), 39-47.
- Luborsky, L., Diguier, L., DeRubeis, R. J., & Schmidt, K. A. (1997). A core battery of measures of depression and principles for their selection. In H. H. Strupp & L. M. Horowitz & M. J. Lambert (Eds.), Measuring Patient Changes in Mood, Anxiety, and Personality Disorders: Toward a core battery (pp. 283-300). Washington, DC: American Psychological Association.
- Luborsky, L., Diguier, L., Seligman, D. A., Rosenthal, R., Krause, E. D., Johnson, S., Halperin, G., Bishop, M., Berman, J. S., & Schweizer, E. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. Clinical Psychology: Science and Practice, 6(1), 95-106.

- Luborsky, L., Mintz, J., Auerback, A., Christoph, P., Bachrach, H., Todd, T., Johnson, M., Cohen, M., & O'Brien, C. P. (1980). Predicting the outcome of psychotherapy: Findings of the Penn Psychotherapy Project. Archives of General Psychiatry, *37*, 471-481.
- Luborsky, L., Popp, C., & Barber, J. P. (1994). Common and special factors in different transference-related measures. Psychotherapy Research, *4*(3&4), 277-286.
- Luborsky, L., Singer, B., & Luborsky, L. (1975). Comparative studies of psychotherapies. Archives of General Psychiatry, *32*, 995-1008.
- Lunnen, K. M., & Ogles, B. M. (1998). A multiperspective, multivariable evaluation of reliable change. Journal of Consulting and Clinical Psychology, *66*(2), 400-410.
- Lyons-Ruth, K. (1996). Attachment relationships among children with aggressive behavior problems: The role of disorganized early attachment patterns. Journal of Consulting and Clinical Psychology, *64*, 64-73.
- Lyons-Ruth, K. (1998). Implicit relational knowing: Its role in development and psychoanalytic treatment. Infant Mental Health Journal, *19*(3), 282-289.
- Mace, C., & Margison, F. (1997). Attachment and psychotherapy: An overview. British Journal of Medical Psychology, *70*, 209-215.
- Main, M. (1991). Recording and transcribing the Adult Attachment Interview. Unpublished manuscript: Department of Psychology, University of California, Berkeley.
- Main, M. (1995). Recent studies in attachment: Overview with selected implications for clinical work. In S. Goldberg & R. Muir & J. Kerr (Eds.), Attachment theory: Social, developmental, and clinical perspectives (pp. 407-474). Hillsdale, NJ: The Analytic Press.
- Main, M. (1996). Introduction to the Special Section on Attachment and Psychopathology: 2. Overview of the field of attachment. Journal of Consulting and Clinical Psychology, *64*(2), 237-243.
- Main, M., & Goldwyn, R. (1985/1994). Adult attachment scoring and classification systems. Unpublished manuscript. Berkeley, CA: University of California.
- Main, M., & Solomon, J. (1990). Procedures for identifying infants as disorganized/disoriented during the Ainsworth Strange Situation. In M. T. Greenberg & D. Cicchetti & E. M.

- Cummings (Eds.), Attachment in the preschool years (pp. 121-160). Chicago: University of Chicago Press.
- Malan, D. H. (1973). The outcome problem in psychotherapy research. Archives of General Psychiatry, *29*, 719-729.
- Mallinckrodt, B. (2000). Attachment, social competencies, social support, and interpersonal process in psychotherapy. Psychotherapy Research, *10*(3), 239-266.
- Manning, W. H., & DuBois, P. H. (1962). Correlational methods in research on human learning. Perceptual and Motor Skills, *15*, 287-321.
- Marmar, C. R., Gaston, L., Gallagher, D., & Thompson, L. W. (1989). Alliance and outcome in late-life depression. Journal of Nervous and Mental Disease, *177*(8), 464-472.
- Marmar, C. R., Weiss, D. S., & Gaston, L. (1989). Toward the validation of the California Therapeutic Alliance Rating System. Psychological Assessment, *1*(1), 46-52.
- Marziali, E. (1984). Three viewpoints on the therapeutic alliance: Similarities, differences, and associations with psychotherapy outcome. Journal of Nervous and Mental Disease, *172*(7), 417-423.
- Marziali, E., Munroe-Blum, H., & McCleary, L. (1999). The effects of the therapeutic alliance on the outcomes of individual and group psychotherapy with borderline personality disorder. Psychotherapy Research, *9*(4), 424-436.
- Mayes, L., & Spence, D. P. (1994). Understanding therapeutic action in the analytic situation: A second look at the developmental metaphor. Journal of the American Psychoanalytic Association, *42*, 789-816.
- McCullough, L., Winston, A., Farber, B. A., Porter, F., Pollack, J., Laikin, M., Vingiano, W., & Trujillo, M. (1991). The relationship of patient-therapist interaction to outcome in brief psychotherapy. Psychotherapy, *28*, 525-533.
- McGlashan, T. H. (1984). The Chestnut Lodge follow-up study: I. Follow-up methodology and study sample. Archives of General Psychiatry, *41*, 573-585.
- McGlashan, T. H., & Miller, G. H. (1982). The goals of psychoanalysis and psychoanalytic psychotherapy. Archives of General Psychiatry, *39*, 377-388.

- McNeilly, C. L., & Howard, K. I. (1991). The Therapeutic Procedures Inventory: Psychometric properties and relationship to phase of treatment. Journal of Psychotherapy Integration, *1*(3), 223-234.
- Mergenthaler, E., & Bucci, W. (1999). Linking verbal and non-verbal representations: computer analysis of referential activity. British Journal of Medical Psychology, *72*(3), 339-354.
- Mergenthaler, E., & Kächele, H. (1996). Applying multiple computerized text-analytic measures to single psychotherapy cases. Journal of Psychotherapy Practice and Research, *5*, 307-317.
- Merten, J., Anstadt, T., Ullrich, B., Krause, R., & Buchheim, P. (1996). Emotional experience and facial behavior during the psychotherapeutic process and its relation to treatment outcome. Psychotherapy Research, *6*(3), 198-212.
- Mickelson, K. D., Kessler, R. C., & Shaver, P. R. (1997). Adult attachment in a nationally representative sample. Journal of Personality and Social Psychology, *73*(5), 1092-1106.
- Milbrath, C., Bond, M., Cooper, S., Znoj, H. J., Horowitz, M. J., & Perry, J. C. (1999). Sequential consequences of therapists' interventions. Journal of Psychotherapy Practice and Research, *8*(1), 40-54.
- Mintz, J. (1981). Measuring outcome in psychodynamic psychotherapy. Archives of General Psychiatry, *38*, 503-506.
- Mintz, J., Luborsky, L., & Christoph, P. (1979). Measuring the outcomes of psychotherapy: Findings of the Penn Psychotherapy Project. Journal of Consulting and Clinical Psychology, *47*(2), 319-334.
- Modell, A. H. (1998). Review of infant mental health papers. Infant Mental Health Journal, *19*(3), 341-345.
- Monsen, J., Odland, T., Faugli, A., Daae, E., & Eilertsen, D. E. (1995a). Personality disorders and psychosocial changes after intensive psychotherapy: A prospective follow-up study of an outpatient psychotherapy project, 5 years after end of treatment. Scandinavian Journal of Psychology, *36*, 256-268.

- Monsen, J., Odland, T., Faugli, A., Daae, E., & Eilertsen, D. E. (1995b). Personality disorders: Changes and stability after intensive psychotherapy focusing on affect consciousness. Psychotherapy Research, *5*(1), 33-48.
- Moran, G., Fonagy, P., Kurtz, A., Bolton, A., & Brook, C. (1991). A controlled study of the psychoanalytic treatment of brittle diabetes. Journal of the American Academy of Child and Adolescent Psychiatry, *30*(6), 926-935.
- Moran, G. S., & Fonagy, P. (1987). Psychoanalysis and diabetic control: A single-case study. British Journal of Medical Psychology, *60*, 357-372.
- Moras, K., & Strupp, H. H. (1982). Pretherapy interpersonal relations, patients' alliance, and outcome in brief therapy. Archives of General Psychiatry, *39*, 405-409.
- Morgan, R., Luborsky, L., Crits-Christoph, P., Curtis, H., & Solomon, J. (1982). Predicting the outcomes of psychotherapy by the Penn Helping Alliance Rating Method. Archives of General Psychiatry, *39*, 397-402.
- Muscetta, S., Dazzi, N., De Coro, A., Ortu, F., & Speranza, A. M. (1999). 'States of mind with respect to attachment' and change in a psychotherapeutic relationship: A study of the coherence of transcript in a short-term psychotherapy with an adolescent. Psychoanalytic Inquiry, *19*(5), 885-921.
- Nahum, J. P. (1998). Case illustration: Moving along . . . and, is change gradual or sudden? Infant Mental Health Journal, *19*(3), 315-319.
- Najavits, L. M. (2001). Helping "difficult" patients. Psychotherapy Research, *11*(2), 131-152.
- Najavits, L. M., & Gunderson, J. G. (1995). Better than expected: Improvements in borderline personality disorder in a 3-year prospective outcome study. Comprehensive Psychiatry, *36*(4), 296-302.
- Nelson, H. E. (1982). National Adult Reading Test (NART), Test Manual. Windsor, Berkshire: NFER-Nelson Publishing.
- Nietzel, M. T., & Trull, T. J. (1988). Meta-analytic approaches to social comparisons: A method for measuring clinical significance. Behavioral Assessment, *10*, 159-169.

- Norman, H., Blacker, K. H., Oremland, J. D., & Barrett, W. G. (1976). The fate of the transference neurosis after termination of a satisfactory analysis. Journal of the American Psychoanalytic Association, *24*, 471-498.
- Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric Theory (Third ed.). New York: McGraw Hill, Inc.
- Oberndorf, C. P. (1943). Results of psycho-analytic therapy. International Journal of Psycho-Analysis, *24*, 107-114.
- O'Connor, J. F., Daniels, G., Karush, A., Moses, L., Flood, C., & Stern, L. O. (1964). The effects of psychotherapy on the course of ulcerative colitis: A preliminary report. American Journal of Psychiatry, *120*, 738-742.
- Ogles, B. M., Lambert, M. J., Weight, D. G., & Payne, I. R. (1990). Agoraphobia outcome measurement: A review and meta-analysis. Psychological Assessment, *2*, 317-325.
- Ogrodniczuk, J. S., & Piper, W. E. (1999). Measuring therapist technique in psychodynamic psychotherapies: development and use of a new scale. Journal of Psychotherapy Practice and Research, *8*, 142-154.
- O'Malley, S. S., Suh, C. S., & Strupp, H. H. (1983). The Vanderbilt Psychotherapy Process Scale: A report on the scale development and a process-outcome study. Journal of Consulting and Clinical Psychology, *51*(4), 581-586.
- OPD Task Force (Ed.). (2001). Operationalized Psychodynamic Diagnostics: Foundations and Manual. Seattle: Hogrefe & Huber.
- Oremland, J. D., Blacker, K. H., & Norman, H. (1975). Incompleteness in "successful" psychoanalyses: A follow-up study. Journal of the American Psychoanalytic Association, *23*, 819-844.
- Orgel, S. Z. (1958). Effects of psychoanalysis on the course of peptic ulcer. Psychosomatic Medicine, *20*, 117-123.
- Orlinsky, D. E., & Geller, J. D. (1993). Patients' representations of their therapists and therapy: New measures. In N. E. Miller & L. Luborsky & J. P. Barber & J. P. Docherty (Eds.),

- Psychodynamic Treatment Research: A Handbok for Clinical Practice (pp. 423-466). New York: Basic Books.
- Orlinsky, D. E., Geller, J. D., Tarragona, M., & Farber, B. (1993). Patients' representations of psychotherapy: A new focus for psychodynamic research. Journal of Consulting and Clinical Psychology, *61*(4), 596-610.
- Orlinsky, D. E., Grawe, K., & Parks, B. K. (1994). Process and outcome in psychotherapy - noch einmal. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of Psychotherapy and Behavior Change (pp. 270-376). New York: John Wiley & Sons, Inc.
- Orlinsky, D. E., & Howard, K. I. (1967). The good therapy hour: Experiential correlates of patients' and therapists' evaluations of therapy sessions. Archives of General Psychiatry, *16*, 621-632.
- Orlinsky, D. E., & Howard, K. I. (1986). The psychological interior of psychotherapy: Explorations with the Therapy Session Reports. In L. S. Greenberg & W. M. Pinsof (Eds.), The psychotherapeutic process: A research handbook (pp. 477-501). New York: Guilford.
- Orlinsky, D. E., Rønnestad, M. H., & Willutzki, U. (2004). Fifty years of psychotherapy process-outcome research: continuity and change. In M. J. Lambert (Ed.), Bergin and Garfield's Handbook of Psychotherapy and Behavior Change (pp. 307-389). New York: John Wiley & Sons, Inc.
- Parloff, M. B. (1986). Placebo controls in psychotherapy research: A sine qua non or a placebo for research problems? Journal of Consulting and Clinical Psychology, *54*(1), 79-87.
- Patrick, M., Hobson, R. P., Castle, D., Howard, R., & Maugan, B. (1994). Personality disorder and the mental representation of early social experience. Development and Psychopathology, *6*, 375-388.
- Paul, G. L. (1967). Strategy of outcome research in psychotherapy. Journal of Consulting Psychology, *31*(2), 109-118.
- Pekarik, G., & Wolff, C. B. (1996). Relationship of satisfaction to symptom change, follow-up adjustment, and clinical significance. Professional Psychology: Research and Practice, *27*, 202-208.

- Perron, R. (1999). Reflections on psychoanalytic research problems - the French speaking view. In IPA (Ed.), An open door review of outcome studies in psychoanalysis (pp. 8-19). London: Research Committee of the International Psychoanalytic Association.
- Perry, J. C. (1989a). Defense mechanism rating scales. Cambridge, MA.
- Perry, J. C. (1989b). Scientific progress in psychodynamic formulation. Psychiatry, *52*, 245-249.
- Perry, J. C. (1994). Assessing psychodynamic patterns using the Idiographic Conflict Formulation Method. Psychotherapy Research, *4*(3&4), 239-252.
- Perry, J. C., Augusto, F., & Cooper, S. H. (1989). Assessing psychodynamic conflicts: I. Reliability of the Idiographic Conflict Formulation Method. Psychiatry, *52*, 289-301.
- Perry, J. C., Luborsky, L., Silberschatz, G., & Popp, C. (1989). An examination of three methods of psychodynamic formulation based on the same videotaped interview. Psychiatry, *52*, 302-323.
- Pessier, J., & Stuart, J. (2000). A new approach to the study of therapeutic work in the transference. Psychotherapy Research, *10*(2), 169-180.
- Pfeffer, A. Z. (1959). A procedure for evaluating the results of psychoanalysis - A preliminary report. Journal of the American Psychoanalytic Association, *7*, 418-444.
- Pianta, R. C., Egeland, B., & Adam, E. K. (1996). Adult attachment classification and self-reported psychiatric symptomatology as assessed by the Minnesota Multiphasic Personality Inventory-2. Journal of Consulting and Clinical Psychology, *64*(2), 273-281.
- Piper, W. E., Azim, H. F. A., Joyce, A. S., McCallum, M., Nixon, G. W. H., & Segal, P. S. (1991). Quality of object relations versus interpersonal functioning as predictors of therapeutic alliance and psychotherapy outcome. Journal of Nervous and Mental Disease, *179*(7), 432-438.
- Piper, W. E., Debbane, W. G., de Carufel, F. L., & Bienvenu, J. P. (1987). A system for differentiating therapist interventions from other interventions. Bulletin of the Menninger Clinic, *51*(6), 532-550.

- Piper, W. E., Joyce, A. S., McCallum, M., & Azim, H. F. A. (1993). Concentration and correspondence of transference interpretations in short-term psychotherapy. Journal of Consulting and Clinical Psychology, *61*(4), 586-595.
- Pole, N., & Jones, E. E. (1998). The talking cure revisited: Content analysis of a two-year psychodynamic psychotherapy. Psychotherapy Research, *8*(2), 171-189.
- Rangell, L. (1954). Similarities and differences between psychoanalysis and dynamic psychotherapy. Journal of the American Psychoanalytic Association, *2*(734-744).
- Rangell, L. (1989). Structural and interstructural change in psychoanalytic treatment. Psychoanalytic Inquiry, *9*(1), 45-66.
- Rapaport, D., & Gill, M. M. (1959). The points of view and assumptions of metapsychology, Collected Papers of David Rapaport (pp. 795-811). New York: Basic Books.
- Research Task Force of the National Institute of Mental Health. (1975). Research in the service of mental health. (DHEW Publication No. ADM 75-236). Rockville, MD.
- Richardson, P. (2001). Evidence-based practice and the psychodynamic psychotherapies. In C. Mace & S. Moorey & B. Roberts (Eds.), Evidence in the Psychological Therapies: A Critical Guide for Practitioners (pp. 157-173). Taylor & Francis: Philadelphia, PA.
- Rosenberg, S. E., Horowitz, L. M., Hanks, S., Hartley, D., Levenson, H., Schulman, t. G., & Skuja, A. (1994). The Consensual Rspnse Psychodynamic Formulation: Part 2: Application to case of Ms. Smithfield. Psychotherapy Research, *4*(3&4), 234-238.
- Rosenfeld, H. (1965). Psychotic states: A psychoanalytic approach. New York: International Universities Press.
- Rosenfeld, H. (1971). A clinical approach to the psychoanalytic theory of the life and death instincts: an investigation into the aggressive aspects of narcissism. International Journal of Psycho-Analysis, *52*, 169-178.
- Rosenstein, D. S., & Horowitz, H. A. (1996). Adolescent attachment and psychopathology. Journal of Consulting and Clinical Psychology, *64*(2), 244-253.
- Rosenzweig, S. (1936). Some implicit common factors in diverse methods in psychotherapy. American Journal of Orthopsychiatry, *6*, 412-415.

- Rosenzweig, S. (1954). A transvaluation of psychotherapy: A reply to Hans Eysenck. Journal of Abnormal and Social Psychology, 49, 298-304.
- Roth, A., & Fonagy, P. (1996). What works for whom: A critical review of psychotherapy research. New York: The Guilford Press.
- Roth, A., & Fonagy, P. (in press). What works for whom: A critical review of psychotherapy research (Second. ed.). New York: Guilford Press.
- Rounsaville, B. J., Chevron, E. S., Prusoff, B. A., Elkin, I., Imber, S., Sotsky, S., & Watkins, J. (1987). The relation between specific and general dimensions of the psychotherapy process in interpersonal psychotherapy of depression. Journal of Consulting and Clinical Psychology, 55(3), 379-384.
- Rubino, G., Barker, C., Roth, T., & Fearon, P. (2000). Therapist empathy and depth of interpretation in response to potential alliance ruptures: The role of therapist and patient attachment styles. Psychotherapy Research, 10(4), 408-420.
- Rudolf, G. (1991). Free University of Berlin: Berlin Psychotherapy Study. In L. E. Beutler & M. Crago (Eds.), Psychotherapy Research: An International Review of Programmatic Studies (pp. 185-193). Washington, D. C.: American Psychological Association.
- Rudolf, G., Grande, T., & Oberbracht, C. (1997, June 19-21). A research strategy for measuring structural change in psychoanalytic psychotherapies. Paper presented at the 20th Ulm Workshop of Empirical Research in Psychoanalysis, Ulm, Germany.
- Russell, R. L., & Trull, T. J. (1986). Sequential analyses of language variables in psychotherapy process research. Journal of Clinical and Consulting Psychology, 54(1), 16-21.
- Safran, J. D., & Aron, L. (2001). Symposium on implications of the Empirically Supported Treatment controversy for psychoanalysis: Introduction. Psychoanalytic Dialogues, 11(4), 571-582.
- Safran, J. D., & Wallner, L. K. (1991). The relative predictive validity of two therapeutic alliance measures in cognitive therapy. Psychological Assessment, 3(2), 188-195.

- Sagi, A., van IJzendoorn, M. H., Scharf, M., Koiren-Karie, M., Joels, T., & Mayseless, O. (1994). Stability and discriminant validity of the Adult Attachment Interview: A psychometric study in young Israeli adults. Developmental Psychology, *30*(5), 771-777.
- Samstag, L. W., Batchelder, S. T., Muran, J. C., Safran, J. D., & Winston, A. (1998). Early identification of treatment failures in short-terms psychotherapy: An assessment of therapeutic alliance and interpersonal behavior. Journal of Psychotherapy Practice and Research, *7*(2), 126-143.
- Sandell, R. (1987a). Assessing the effects of psychotherapy: I. Analysis and critique of present conventions of estimating change. Psychotherapy and Psychosomatics, *47*, 29-36.
- Sandell, R. (1987b). Assessing the effects of psychotherapy: II. A procedure for direct rating of therapeutic change. Psychotherapy and Psychosomatics, *47*, 37-43.
- Sandell, R. (1987c). Assessing the effects of psychotherapy: III. Reliability and validity of 'Change after Psychotherapy'. Psychotherapy and Psychosomatics, *47*, 44-52.
- Sandell, R. (1988). Our varying ability to predict the outcomes of psychotherapy. Psychotherapy and Psychosomatics, *50*, 134-140.
- Sandell, R. (1991). A closer look at the ability to predict psychotherapeutic outcome. Counselling Psychology Quarterly, *4*(2/3), 127-134.
- Sandell, R., Blomberg, J., Lazar, A., Carlsson, J., Broberg, J., & Schubert, J. (2000). Varieties of long-term outcome among patients in psychoanalysis and long-term psychotherapy. International Journal of Psychoanalysis, *81*, 921-942.
- Sandler, J., & Dreher, A. U. (1996). What do psychoanalysts want? The problem of aims in psychoanalytic therapy. London: Routledge.
- Sandler, J., Kennedy, H., & Tyson, R. L. (1980). The technique of child psychoanalysis: Discussions with Anna Freud. London: Karnac Books.
- Sashin, J. I., Eldred, S. H., & Van Amerongen, S. T. (1975). A search for predictive factors in institute supervised cases: A retrospective study of 183 cases from 1959-1966 at the Boston psychoanalytic society and institute. International Journal of Psycho-analysis, *56*, 343-360.

- Sauer, E. M., Lopez, F. G., & Gormley, B. (2003). Respective contributions of therapist and client adult attachment orientations to the development of early working alliance: A preliminary growth modeling study. Psychotherapy Research, 13(3), 371-382.
- Saunders, S. M. (1999). Clients' assessment of the affective environment of the psychotherapy session: Relationship to sessional quality and treatment effectiveness. Journal of Clinical Psychology, 55(5), 597-605.
- Schacht, T. E., & Henry, W. P. (1994). Modeling recurrent patterns of interpersonal relationship with Structural Analysis of Social Behavior: The SASB-CMP. Psychotherapy Research, 4(3&4), 208-221.
- Scharf, R. D., Waldron, S. J., Firestein, S. K., Goldberger, A., & Burton, A. (1999). The analytic process scales (APS) coding manual: Unpublished manuscript.
- Schjelderup, H. (1955). Lasting effects of psychoanalytic treatment. Psychiatry, 18, 103-133.
- Schlesinger, H. J. (1974). Problems of doing research on the therapeutic process in psychoanalysis. Journal of the American Psychoanalytic Association, 22, 3-13.
- Schlessinger, N., & Robbins, F. (1974). Assessment and follow-up in psychoanalysis. Journal of the American Psychoanalytic Association, 22, 542-567.
- Schwartz, F. (1981). Psychic structure. International Journal of Psycho-Analysis, 62, 61-72.
- Scott Brown, L., & Wright, J. (2003). The relationships between attachment strategies and psychopathology in adolescence. Psychology and Psychotherapy: Theory, Research, and Practice, 76, 351-367.
- Sechrest, L. (1994). Recipes for psychotherapy. Journal of Clinical and Consulting Psychology, 62(5), 952-954.
- Segal, Z. V. (1997). Implications of priming for measures of change following psychological and pharmacological treatments. In H. H. Strupp & L. M. Horowitz & M. J. Lambert (Eds.), Measuring Patient Changes in Mood, Anxiety, and Personality Disorders: Toward a core battery (pp. 81-98). Washington, DC: American Psychological Association.

- Seidenstrücker, G., & Baumann, U. (1978). Multimethodale diagnostik. In U. Baumann & H. Berbalk & G. Seidenstrücker (Eds.), Klinische Psychologie: Trends in Forschung und Praxis (Vol. 1, pp. 134-182). Bern, Switzerland: Huber.
- Semrad, E. V., Grinspoon, L., & Fienberg, S. E. (1973). Development of an ego profile scale. Archives of General Psychiatry, *28*, 70-77.
- Senf, W., Kordy, H., von Rad, M., & Bräutigam, W. (1984). Indication in psychotherapy on the basis of a follow-up study. Psychotherapy and Psychosomatics, *42*, 37-47.
- Sexton, H. (1993). Exploring a psychotherapeutic change sequence: Relating process to intersessional and posttreatment outcome. Journal of Consulting and Clinical Psychology, *61*(1), 128-136.
- Shadish, W. R. J., Matt, G. E., Navarro, A. M., Siegle, G., Crits-Christoph, P., Hazelrigg, M. D., Jorm, A. F., Lyons, L. C., Nietzel, M. T., Prout, H. T., Robinson, L., Smith, M. L., Svartberg, M., & Weiss, B. (1997). Evidence that treatment works in clinically representative conditions. Journal of Consulting and Clinical Psychology, *65*(3), 355-365.
- Shapiro, D. A., & Shapiro, D. (1982). Meta-analysis of comparative therapy outcome studies: A replication and refinement. Psychological Bulletin, *92*(3), 581-604.
- Shapiro, T. (Ed.). (1991). The concept of structure in psychoanalysis. Madison, CT: International Universities Press, Inc.
- Siegel, N. H. (1962). Characteristics of patients in psychoanalysis. Journal of Nervous and Mental Disease, *135*, 155-158.
- Sifneos, P. E., Apfel, R. J., Bassuk, E., Fishman, G., & Gill, A. (1980). Ongoing outcome research on short-term dynamic psychotherapy. Psychotherapy and Psychosomatics, *33*, 233-241.
- Silberschatz, G. (1994). Abuse and disabuse of the drug metaphor in psychotherapy research: Hold on to the baby as you throw out the bath. Journal of Consulting and Clinical Psychology, *62*(5), 949-951.
- Silberschatz, G., Fretter, P. B., & Curtis, J. T. (1986). How do interpretations influence the process of psychotherapy? Journal of Consulting and Clinical Psychology, *54*, 646-652.

- Slade, A. (1999). Attachment theory and research: Implications for the theory and practice of individual psychotherapy with adults. In J. Cassidy & P. R. Shaver (Eds.), Handbook of attachment: Theory, research, and clinical applications (pp. 575-594). New York: Guilford Press.
- Smith, M., Glass, G., & Miller, T. (1980). The Benefits of Psychotherapy. Baltimore, MD: Johns Hopkins University Press.
- Smith, S. R., Hilsenroth, M. J., Baity, M. R., & Knowles, E. S. (2003). Assessment of patient and therapist perspectives of process. American Journal of Psychotherapy, 57(2), 195-205.
- Snow, R. E. (1991). Aptitude-treatment interaction as a framework for research on individual differences in psychotherapy. Journal of Consulting and Clinical Psychology, 59(2), 205-216.
- Soloff, P. H., & Ulrich, R. F. (1981). Diagnostic Interview for Borderline patients: A replication study. Archives of General Psychiatry, 38, 686-692.
- Spangler, G., & Grossmann, K. E. (1993). Biobehavioral organization in securely and insecurely attached infants. Child Development, 64, 1439-1450.
- Speer, D. C. (1992). Clinically significant change: Jacobson and Truax (1991) revisited. Journal of Consulting and Clinical Psychology, 60(3), 402-408.
- Speer, D. C. (1999). What is the role of two-wave designs in clinical research? Comment on Hageman and Arrindell. Behaviour Research and Therapy, 37, 1203-1210.
- Speer, D. C., & Greenbaum, P. E. (1995). Five methods for computing significant individual client change and improvement rates: Support for an individual growth curve approach. Journal of Consulting and Clinical Psychology, 63(6), 1044-1048.
- Spence, D., Mayes, L. C., & Dahl, H. (1994). Monitoring the analytic surface. Journal of the American Psychoanalytic Association, 42, 43-64.
- Spence, D. P. (1998). Pronoun co-occurrence as a measure of shared understanding. In R. F. Bornstein & J. M. Masling (Eds.), Empirical Studies of the Therapeutic Hour (pp. 153-178). Washington, DC: American Psychological Association.

- Spence, D. P., Dahl, H., & Jones, E. E. (1993). Impact of interpretation on associative freedom. Journal of Consulting and Clinical Psychology, *61*(3), 395-402.
- Spielberger, C. D. (1983). Manual for the State-Trait Anxiety Inventory, STAI (Form Y) ("Self-Evaluation Questionnaire"). Palo Alto, CA: Consulting Psychologists Press.
- Spitzer, R. L., Endicott, J., & Robins, E. (1978). Research diagnostic criteria: Rationale and reliability. Archives of General Psychiatry, *35*, 773-782.
- Statistical Sciences. (1995). S-PLUS User's Manual, Version 3.3 for Windows. Seattle: Statistical Sciences, Inc.
- Steimer-Krause, E., Krause, R., & Wagner, G. (1990). Interaction regulations used by schizophrenic and psychosomatic patients. Psychiatry, *53*, 209-228.
- Stein, H., Jacobs, N., Ferguson, K., Allen, J. G., & Fonagy, P. (1998). What do adult attachment scales measure? Bulletin of the Menninger Clinic, *62*(1), 33-82.
- Stern, D. N. (1998). The process of therapeutic change involving implicit knowledge: Some implications of developmental observations for adult psychotherapy. Infant Mental Health Journal, *19*(3), 300-308.
- Stiles, W. B. (1980). Measurement of the impact of psychotherapy sessions. Journal of Consulting and Clinical Psychology, *48*(2), 176-185.
- Stiles, W. B. (1986). Development of a taxonomy of Verbal Response Modes. In L. S. Greenberg & W. M. Pinsof (Eds.), The psychotherapeutic process: A research handbook (pp. 161-199). New York: Guilford.
- Stiles, W. B. (1994). Drugs, recipes, babies, bathwater, and psychotherapy process-outcome relations. Journal of Clinical and Consulting Psychology, *62*(5), 955-959.
- Stiles, W. B., Elliott, R., Llewelyn, S. P., Firth-Cozens, J., Margison, F. R., Shapiro, D. A., & Hardy, G. (1990). Assimilation of problematic experiences by clients in psychotherapy. Psychotherapy, *27*(3), 411-420.
- Stiles, W. B., Meshot, C. M., Anderson, T. M., & Sloan, W. W. J. (1992). Assimilation of problematic experiences: The case of John Jones. Psychotherapy Research, *2*(2), 81-101.

- Stiles, W. B., Morrison, L. A., Haw, S. K., Harper, H., Shapiro, D. A., & Firth-Cozens, J. (1991). Longitudinal study of assimilation in exploratory psychotherapy. Psychotherapy, *28*(2), 195-206.
- Stiles, W. B., & Shapiro, D. A. (1989). Abuse of the drug metaphor in psychotherapy process-outcome research. Clinical Psychology Review, *9*, 521-543.
- Stiles, W. B., & Shapiro, D. A. (1994). Disabuse of the drug metaphor: Psychotherapy process-outcome correlations. Journal of Consulting and Clinical Psychology, *62*(5), 942-948.
- Stolorow, R. D. (1978). The concept of psychic structure: Its metapsychological and clinical psychoanalytic meanings. International Review of Psycho-analysis, *5*, 313-320.
- Stone, A. A., Turkkan, J. S., Bachrach, C. A., Jobe, J. B., Kurtzman, H. S., & Cain, V. S. (Eds.). (2000). The Science of Self-report. Mahwah, New Jersey: Lawrence Erlbaum.
- Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. K. Denzin & Y. S. Lincoln (Eds.), Handbook of Qualitative Research (pp. 273-285). Thousand Oaks, California: Sage.
- Strupp, H. H. (1986). Psychotherapy: Research, practice and public policy (how to avoid dead ends). American Psychologist, *41*(2), 120-130.
- Strupp, H. H. (2001). Implications of the Empirically Supported Treatment movement for psychoanalysis. Psychoanalytic Dialogues, *11*(4), 605-620.
- Strupp, H. H., & Hadley, S. W. (1979). Specific versus nonspecific factors in psychotherapy: A controlled study of outcome. Archives of General Psychiatry, *36*, 1125-1136.
- Strupp, H. H., Horowitz, L. M., & Lambert, M. J. (Eds.). (1997). Measuring Patient Changes in Mood, Anxiety, and Personality Disorders. Washington, DC: American Psychological Association.
- Strupp, H. H., Schacht, T. E., & Henry, W. P. (1988). Problem-Treatment-Outcome congruence: A principle whose time has come. In H. Dahl & H. Kächele & H. Thomä (Eds.), Psychoanalytic process research strategies (pp. 1-14). Berlin: Springer-Verlag.
- Suh, C. S., Strupp, H. H., & O'Malley, S. S. (1986). The Vanderbilt Process Measures: The Psychotherapy Process Scale (VPPS) and the Negative Indicators Scale (VNIS). In L. S.

- Greenberg & W. M. Pinsof (Eds.), The psychotherapeutic process: A research handbook (pp. 285-323). New York: The Guilford Press.
- Sundin, E., & Armelius, B.-Å. (1998). Mental health and psychic structure: An empirical study. Bulletin of the Menninger Clinic, *62*(4), 475-489.
- Tabachnik, B. G., & Fidell, L. S. (1989). Using Multivariate Statistics. New York: HarperCollins.
- Thorngate, W. (1986). The production, detection, and explanation of behavioral patterns. In J. Valsiner (Ed.), The Individual Subject and Scientific Psychology (pp. 71-93). New York: Plenum Press.
- Tichenor, V., & Hill, C. E. (1989). A comparison of six measures of working alliance. Psychotherapy, *26*(2), 195-199.
- Tingey, R. C., Lambert, M. J., Burlingame, G. M., & Hansen, N. B. (1999). Assessing clinical significance: Proposed extensions to method. Psychotherapy Research, *6*(2), 109-123.
- Tracey, T. J., & Kokotovic, A. M. (1989). Factor structure of the Working Alliance Inventory. Psychological Assessment, *1*(3), 207-210.
- Travis, L. A., Binder, J. L., Liwise, N. G., & Horne-Moyer, H. L. (2001). Changes in clients' attachment styles over the course of time-limited dynamic psychotherapy. Psychotherapy: Theory, Research, and Practice, *38*(2), 149-159.
- Tronick, E. Z. (1998a). Dyadically expanded states of consciousness and the process of therapeutic change. Infant Mental Health Journal, *19*(3), 290-299.
- Tronick, E. Z. (1998b). Interventions that effect change in psychotherapy: A model based on infant research. Infant Mental Health Journal, *19*(3), 277-279.
- Tuckett, D. (1993). Some thoughts on the presentation and discussion of the clinical material of psychoanalysis. International Journal of Psycho-analysis, *74*, 1175-1189.
- Tuckett, D. (1994). Developing a grounded hypothesis to understand a clinical process: The role of conceptualisation in validation. International Journal of Psycho-analysis, *75*, 1159-1180.

- Turton, P., McGauley, G., Marin-Avellan, L., & Hughes, P. (2001). The Adult Attachment Interview: Rating and classification problems posed by non-normative samples. Attachment and Human Development, 3(3), 284-303.
- Tyrell, C., & Dozier, M. (1997). The role of attachment in therapeutic process and outcome for adults with serious psychiatric disorders. Washington, DC: Paper presented at the biennial meeting of Society for Research in Child Development.
- Tyrell, C. L., Dozier, M., Teague, G. B., & FalLOT, R. D. (1999). Effective treatment relationships of persons with serious psychiatric disorders: The importance of attachment states of mind. Journal of Consulting and Clinical Psychology, 57(5), 725-733.
- Tyson, P., & Morris, J. L. (1992). Psychoanalysis and psychoanalytic psychotherapy -- similarities and differences: Therapeutic technique. Journal of the American Psychoanalytic Association, 40, 211-221.
- Vaillant, G. E. (1971). Theoretical hierarchy of adaptive ego mechanisms. Archives of General Psychiatry, 24, 107-118.
- Vaillant, G. E. (1993). The wisdom of the ego. Cambridge, MA: Harvard University Press.
- van IJzendoorn, M. H. (1995a). Adult attachment representations, parental responsiveness, and infant attachment: A meta-analysis on the predictive validity of the Adult Attachment Interview. Psychological Bulletin, 117(3), 387-403.
- van IJzendoorn, M. H. (1995b). Of the way we are: On temperament, attachment, and the transmission gap: A rejoinder to Fox (1995). Psychological Bulletin, 117(3), 411-415.
- van IJzendoorn, M. H., & Bakermans-Kranenburg, M. J. (1996). Attachment representations in mothers, fathers, adolescents, and clinical groups: A meta-analytic search for normative data. Journal of Consulting and Clinical Psychology, 64(1), 5-7.
- Vaughn, S. C., & Roose, S. P. (1995). The analytic process: Clinical and research definitions. International Journal of Psychoanalysis, 76, 343-356.
- Waelder, R. (1963). Psychic determinism and the possibility of predictions. Psychoanalytic Quarterly, 32, 15-42.

- Waldron, S., Jr., Scharf, R. D., & Firestein, S. K. (1997). Measuring psychoanalytic work and benefit: The analytic process scales: Unpublished manuscript.
- Waldron, S. J., Scharf, R. D., Hurst, D., Firestein, S. K., & Burton, A. (2004). What happens in a psychoanalysis? A view through the lens of the analytic process scales (APS). International Journal of Psychoanalysis, 85(2), 443-466.
- Wallerstein, R. S. (1968). Discussion of Engel 1968. Journal of the American Psychoanalytic Association, 16, 215-223.
- Wallerstein, R. S. (1986). Forty-Two Lives in Treatment: A study of psychoanalysis and psychotherapy. New York: The Guilford Press.
- Wallerstein, R. S. (1988). Assessment of structural change in psychoanalytic therapy and research. Journal of the American Psychoanalytic Association, 36 (Suppl.), 241-261.
- Wallerstein, R. S. (1995). The talking cures: the psychoanalyses and the psychotherapies. New Haven, CT: Yale University Press.
- Wallerstein, R. S. (2001). Psychotherapy research and psychoanalytic practice: Commentary on papers by Lester Luborsky and Hans H. Strupp. Psychoanalytic Dialogues, 11(4), 621-634.
- Wallerstein, R. S., & DeWitt, K. N. (1997). Intervention modes in psychoanalysis and in psychoanalytic psychotherapies: A revised classification. Journal of Psychotherapy Integration, 7(2), 129-150.
- Wallerstein, R. S., & Sampson, H. (1971). Issues in research in psychoanalytic process. International Journal of Psycho-analysis, 52, 11-49.
- Wampold, B. E. (2001). The Great Psychotherapy Debate: Models, methods, and findings. Mahwah, New Jersey: Lawrence Erlbaum.
- Wampold, B. E., & Jenson, W. R. (1986). Clinical significance revisited [Letter to the editor]. Behavior Therapy, 17, 302-305.
- Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H.-n. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, "All must have prizes". Psychological Bulletin, 122(3), 203-215.

- Warren, F., & Norton, K. (2004). Henderson Hospital democratic therapeutic community: outcome studies and methodological issues. In P. Richardson & H. Kächele & C. Renlund (Eds.), Research on Psychoanalytic Psychotherapy with Adults (pp. 133-153). London: Karnac.
- Warren, S. L., Huston, L., Egeland, B., & Sroufe, L. A. (1997). Child and adolescent anxiety disorders and early attachment. Journal of the American Academy of Child and Adolescent Psychiatry, *36*(5), 637-644.
- Waskow, I. E., & Parloff, M. B. (Eds.). (1975). Psychotherapy change measures (DHEW Publication No. ADM 74-120). Washington, DC: US Government Printing Office.
- Weber, J. J., Bachrach, H. M., & Solomon, M. (1985a). Factors associated with the outcome of psychoanalysis: Report of the Columbia Psychoanalytic Center Research Project (II). International Review of Psychoanalysis, *12*, 127-141.
- Weber, J. J., Bachrach, H. M., & Solomon, M. (1985b). Factors associated with the outcome of psychoanalysis: Report of the Columbia Psychoanalytic Center Research Project (III). International Review of Psychoanalysis, *12*, 251-262.
- Weber, J. J., Solomon, M., & Bachrach, H. M. (1985). Characteristics of psychoanalytic clinic patients: Report of the Columbia Psychoanalytic Center Research Project (I). International Review of Psychoanalysis, *12*, 13-26.
- Weinryb, R. M., Barber, J. P., Foltz, C., Göransson, S. G. M., & Gustavsson, J. P. (2000). The Central Relationship Questionnaire (CRQ): Psychometric properties in a Swedish sample and cross-cultural studies. Journal of Psychotherapy Practice and Research, *9*(4), 201-212.
- Weinryb, R. M., Rössel, R., & Åsberg, M. (1991a). The Karolinska Psychodynamic Profile. I. Validity and dimensionality. Acta Psychiatrica Scandinavica, *83*, 64-72.
- Weinryb, R. M., Rössel, R., & Åsberg, M. (1991b). The Karolinska Psychodynamic Profile. II. Interdisciplinary and cross-cultural reliability. Acta Psychiatrica Scandinavica, *83*, 73-76.
- Weintraub, W., & Aronson, H. (1968). A survey of patients in classical psychoanalysis: Some vital statistics. Journal of Nervous and Mental Disease, *146*, 98-102.

- Weiss, I., Rabinowitz, J., & Spiro, S. (1996). Agreement between therapists and clients in evaluating therapy and its outcomes: Literature review. Administration and Policy in Mental Health, *23*(6), 493-511.
- Weiss, J. (2003). Development of a research program. Psychoanalytic Inquiry, *23*(2), 350-366.
- Weiss, J., & Sampson, H. (1986). The psychoanalytic process: Theory, clinical observations, and empirical research. New York: Guilford.
- Weiss, R. D., Najavits, L. M., Muenz, L. R., & Hufford, C. (1995). Twelve-month test-retest reliability of the Structured Clinical Interview for DSM-III-R personality disorders in cocaine-dependent patients. Comprehensive Psychiatry, *36*(5), 384-389.
- Weissman, M. M., Prusoff, B. A., & Paykel, E. S. (1972). Checklist quantification of a psychological therapy: Pilot studies of reliability and utility. Journal of Nervous and Mental Disease, *154*(2), 125-136.
- Weissman, M. M., Prusoff, B. A., Thompson, W. D., Harding, P. S., & Myers, J. K. (1978). Social adjustment by self-report in a community sample and in psychiatric outpatients. Journal of Nervous and Mental Disease, *166*(5), 317-326.
- Werman, D. S. (1989). The idealization of structural change. Psychoanalytic Inquiry, *9*(1), 119-139.
- Westen, D., Feit, A., & Zittel, C. (1999). Methodological issues in research using projective techniques. In P. C. Kendall & J. N. Butcher & G. Holmbeck (Eds.), Handbook of research methods in clinical psychology (2nd ed., pp. 224-240). New York: Wiley.
- Westen, D., & Morrison, K. (2001). A multidimensional meta-analysis of treatments for depression, panic, and generalized anxiety disorder: An empirical examination of the status of empirically supported therapies. Journal of Consulting and Clinical Psychology, *69*(6), 875-899.
- Wilczek, A., Weinryb, R. M., Barber, J. P., Gustavsson, J. P., & Åsberg, M. (2000). The Core Conflictual Relationship Theme (CCRT) and psychopathology in patients selected for dynamic psychotherapy. Psychotherapy Research, *10*(1), 100-113.

- Wilke, S. (1997, June 19-21). Qualitative analysis of therapists' process notes. Paper presented at the 20th Ulm Workshop of Empirical Research in Psychoanalysis, Ulm, Germany.
- Williams, J. B. W., Gibbon, M., First, M. B., Spitzer, R. L., Davies, M., Borus, J., Howes, M. J., Kane, J., Pope, H. G., Jr., Rounsaville, B., & Wittchen, H.-U. (1992). The Structured Clinical Interview for DSM-III-R (SCID): II. Multisite test-retest reliability. Archives of General Psychiatry, *49*, 630-636.
- Windholz, M. J., & Silberschatz, G. (1988). Vanderbilt psychotherapy process scale: A replication with adult outpatients. Journal of Clinical and Consulting Psychology, *56*(1), 56-60.
- Winnicott, D. W. (1962). The aims of psychoanalytic treatment, The Maturation Processes and the Facilitating Environment (pp. 166-170). London: Hogarth Press.
- Zetzel, E. (1968). The so-called good hysteric, The Capacity for Emotional Growth (pp. 128-142). New York: International Universities Press.
- Zetzel, E. R. (1968). Discussion of Engel 1968. Journal of the American Psychoanalytic Association, *16*, 223-227.
- Zilberg, N. J., Wallerstein, R. S., Dewitt, K. N., Hartley, D., & Rosenberg, S. E. (1991). A conceptual analysis and strategy for assessing structural change. Psychoanalysis and Contemporary Thought, *14*, 317-342.
- Zuckerman, D. M., Prusoff, B. A., Weissman, M. M., & Padian, N. S. (1980). Personality as a predictor of psychotherapy and pharmacotherapy outcome for depressed outpatients. Journal of Consulting and Clinical Psychology, *48*(6), 730-735.