Richardson P, Kächele H, Rendlund C

Research on Psychoanalytic Psychotherapy with Adults

Karnac. London

2003

CONTENTS

Introduction

The present volume is the seventh in the EFPP Karnac Clinical Monograph Series, the first of which appeared in 1995. It is the first devoted specifically to research in adult psychoanalytic psychotherapy and represents a collaborative effort by editors and authors from Finland, Germany, Sweden and the United Kingdom. As such it reflects the global aims of the European Federation of Psychoanalytic Psychotherapy in the Public Sector in contributing to the development of a pan-European community of psychoanalytic psychotherapists. Further monographs are planned under the overall guiding editorship of Professor John Tsiantis, the Chief Editor of the Series. These will focus on psychoanalytic psychotherapy research with children and on therapist issues in the processes and outcomes of psychoanalytic psychotherapy. Interested readers should visit the EFPP website at www.efpp.org

In the present-day culture of evidence based practice as a guiding principle for the delivery of public and private sector health services the critical importance of collating empirical research findings relating to psychoanalytic psychotherapy cannot be overstated. Evidence based clinical guidelines are increasingly finding their way into the mental health arena (see for example the UK Department of Health Treatment Choice Guideline for Psychological Therapies and Counselling developed for general medical practitioners and specialists working in the British National Health Service ñ Parry 2001) and, as of yet, the place of psychoanalytic psychotherapy within such guidelines is far from extensive. The present monograph brings together a number of research reports and overviews all of which have used conventional empirical research methodologies and illustrate, we believe, the potential of such methods to explore questions of real significance to psychoanalytic psychotherapists throughout Europe. Peter Fonagyís excellent Foreword locates the studies within an overview of the contemporary research context.

We are indebted to the series editors John Tsiantis (Chief Editor) and Brian Martindale (Associate Editor for the Individual Psychoanalytic Psychotherapy Section of EFPP) for their patient and helpful guidance in the production of this monograph. We also wish to thank Amaryllis Holland for her editorial acumen in the proof reading of the individual papers, as well as Beverley Foster-Davis for additional editorial support.

Phil Richardson, Horst Kaechele and Camilla Renlund
Editors

Foreword

Psychoanalysis has not fared well in the era of evidence-based medicine (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996). While the United Kingdom and all countries of continental Europe are forced to rationalise healthcare costs to a greater or lesser extent, few general principles have emerged to offer reasonable and ethical grounds for such rationalisation (Knapp, 1997). Arguably, as a consequence, the availability of sound empirical research findings to support the provision of particular treatments has become a requirement of statutory funding, in psychiatry as in other specialties. There can be no doubt that the provision of services on the basis of evidence of effectiveness is preferable to service distribution based on postcodes or luck. Nevertheless, there is quite a high price to pay for equity.

Not surprisingly, expensive, long-term treatments like psychoanalytic therapy have quickly fallen victim to EBM-orientated criteria (Westen & Morrison, in preparation). Evidence for the effectiveness of psychoanalytic therapy is lacking. Why? Short-term treatments are undoubtedly easier to research using random assignments for both practical and ethical reasons. At a more impressionistic level, we might say that the world-view that is normally created by working intensively and on a long-term basis with individuals suffering from relatively enduring and severe mental disorders is incompatible with the ethos of tightly controlled investigations of efficacy. Those who work at close quarters with the human mind will inevitably have an impression of reductionism when they see the full complexity of an individual's struggle with internal and external experience reduced to a single 100-point scale (Endicott, Spitzer, Heiss, & Cohen, 1976; Shaffer et al., 1983) or even 12 five point ones (Wing, Curtis, & Beevor, 1996; Wing, Lelliott, & Beevor, 2000). On top of that, even the anchor points are badly defined. Moreover, the spirit of late 20th Century pragmatism and utilitarianism, perhaps by analogy with the remarkable technological advances of this period, often equated the novel with the good and the traditional with the outdated (Giddens, 1999). The rapid progress of technology and biological science has held out the possibility of biochemical rehabilitation, which many continue to feel offers the only viable solution to the challenge of treating mental disorder. This is already a toxic mixture for psychoanalytic thinking, but add the history of unashamed arrogance of many of our psychodynamic colleagues, who until recently have all too frequently treated their non-psychoanalytically trained mental health worker colleagues with at best benevolent tolerance and at worst contempt and disdain, and you have the

complete background to the current crisis for psychoanalytically orientated psychotherapies.

To take but one example from the US where the crisis was first to come to a head: "Although long-term, so-called intensive therapy has been dying for years, some of our profession's leaders cling fiercely to the illusion that it works and that only psychiatrists can do it. However, since we have proof only of its high cost and not its effectiveness, psychiatry's reluctance to admit that the emperor is indeed naked only increases public scepticism" (Detre & McDonald, 1997, p. 203). The hostility towards psychoanalytic ideas that currently dominates the United States may have its historical origins in the fierce competition for power and control between biological psychiatry and psychoanalysis in US medical schools (Cooper, 1996; Michels, 1994). It is hard to envisage such a titanic struggle in a European context where psychoanalysis has never fully dominated the healthcare system. Limited versions of the same conflicts are, however, evident (Chiesa & Fonagy, 1999). While half a century ago psychoanalysis could credibly present itself as the sole form of humane mental health care (Menninger, Mayman, & Pruyser, 1963), current alternatives to psychoanalytic therapy are mostly relatively sophisticated, well-structured and by no means mindless interventions. Cognitive-behaviour therapy and psycho-pharmacological treatments have powerful effects and are reasonably well tolerated by users.

Across Europe, expert groups in many countries are busy surveying the literature in sincere attempts to identify the treatments that may be most helpful for their citizens who suffer from enduring mental disorder (e.g. Department of Health, 1995; Health Council of the Netherlands, 2001; Weisz, Hawley, Pilkonis, Woody, & Follette, 2000). Statutory funding for psychological therapy is threatened in many countries by the readiness with which pharmacological treatments can be made available to relatively large groups. Popular views concerning the causes of mental illness have changed during 'the decade of the brain', in many places powerfully supported by far-sighted pharmaceutical companies, to the point where commonly held theories of psychological disorder have shifted towards the constitutional and antidepressants are bizarrely accepted as appropriate means of addressing social difficulties (Cornwell, 1996). Behaviour genetics research has not helped (Fonagy, in press). The limited range of environments sampled by most studies and the tendency to conflate error variance with non-shared environment have combined to undermine psychodynamic claims concerning the causal significance of shared early family environment, the bread and butter of psychotherapeutic narrative (Rutter, 2000).

All this is not to say that every recent development has been inconsistent with a psychoanalytic understanding. For example, few now believe that severe psychological disorders are episodic conditions that can be addressed in the long term by a short-term intervention. Problems persist after brief interventions have done their best (e.g. Shea et al., 1992). There has been a backlash against the reification of findings from randomised controlled trials, particularly the limitation on the generalisability of findings emerging from very tightly controlled investigations (Markowitz & Street, 1999; Weisz & Jensen, 1999). There has been a heartfelt outcry for effectiveness rather than efficacy research, the former ostensibly creating a more truthful representation of the value of a treatment in the field (Wells, 1999). There has been a resurgence of interest in qualitative as opposed to quantitative data gathering (e.g. Mayes & Pope, 2000) and so-called expert groups who have generated prescriptive lists of approved therapies have come under occasionally severe criticism (Weisz et al., 2000). Sometimes brain research has been successfully introduced to advance the psychodynamic cause, with studies providing striking support for classical ideas receiving quite extensive coverage (Solms, 2000; Solms & Nersessian, 1999). Perhaps even more important have been related social initiatives based on the assumption of 'developmental programming', early influences bringing about enduring change in neural structure and function (Hertzman & Wiens, 1996). The empirical literature on the long-term and trans-generational effects of quality of parent-infant attachment has also helped to make psychoanalytic concerns with infancy more plausible (O'Connor, Bredenkamp, Rutter, & team, 1999; O'Connor, Rutter, & team, in press). User-led research exploring the strategies of individuals living and coping with mental distress also stresses familiar dynamic themes such as the importance of relationships with family and friends, self-esteem drawn from peer groups and respectful treatment by professionals (Faulkner, 2000; Rose, 2001). Nonetheless, these and other developments have done little to reverse the underlying trend away from long-term psychodynamic therapies towards long-term pharmacological or short-term psychological (usually cognitive-behavioural) treatments.

To counteract these trends, three developments to the currently dominant knowledge base will be necessary. First, we require evidence concerning the specific patient groups who uniquely benefit from psychoanalytic interventions and related to this, assessment systems that help to identify these individuals, either in terms of diagnosis and symptomatology, or in terms of characteristic modes of mental functioning or even social conditions. Second, we need

sensitive measurement systems that identify changes in psychological functioning associated with long-term psychoanalytic therapy that may go beyond symptomatic improvement and indicate benefits that are either valued by clients (or carers) or can be shown to be predictive of relative freedom from future difficulties (prevention). Third, we need to develop new adaptations of psychoanalytic therapy that extend and improve upon existing applications to increase their generalisability across clinical groups and enhance their impact either in terms of symptom relief or prevention.

The contributors to the present volume have all, in their own ways, advanced one or more of the above goals. An excellent well-controlled study attempting to address the first of these three aims comes from Horst Kächele's laboratory in Stuttgart. Kächele is a key figure in empirical psychoanalysis, bringing infectious enthusiasm mixed with an iconoclastic approach to his subject. The large-scale eating disorder study attempted to identify which patients with eating disorder problems were particularly well suited to a psychoanalytic approach in in-patient treatment. Whilst a clearly identifiable group benefited remarkably from the treatment they were offered, their success could not be predicted by a traditional questionnaire measure of personality. Perhaps a new approach to measurement of individual differences is required by psychoanalytic investigations.

A suggestion about a possible direction for such a project is provided by the programme of work on in-patient psychotherapy reported from the University of Jena. Bernhard Strauss and colleagues looked at the Inventory of Interpersonal Problems (IIP) as a predictor of therapeutic benefit. They report that psychodynamic treatment may be, somewhat paradoxically, most appropriate for patients who report a wider variety and severity of interpersonal problems. This may be linked to the fundamentally interpersonal focus of modern psychoanalytic therapy. The greater benefit of those with high IIP scores is most likely accounted for by the assumption that those who have greatest awareness of their interpersonal problems are most likely to benefit from a treatment that has interpersonal understanding as its core aim. It is hard to know from a psychodynamic standpoint how responses from a self-report questionnaire on relationship problems might interface with the social behaviour of individuals in an in-patient group therapy setting. It seems likely that only the combination of observational and self-report measures will ultimately yield conclusive answers concerning the precise relationship of initial self-awareness and overt behaviour as predictors of responsiveness to an insight orientated therapy.

The Heidelberg-Berlin study is an excellent example of an attempt to show that long-term intensive therapy may yield added treatment value for severe psychological difficulties when the measurement system is up to the task of showing key differences in character, including relationship patterns, conflict types and structural capacities. The OPD system described in this chapter has become an enormously valuable addition to the empirical armamentarium of psychoanalytic clinicians (Cierpka et al., 1995). A key limitation of the field that has hindered the cumulative construction of a psychoanalytic knowledge base has been the absence of even a rudimentary classification system to describe clinical cases (Gabbard, Gunderson, & Fonagy, in press). A similar initiative with equally great potential is the Munich Psychotherapy Project offering a thorough test of the Scale of Psychological Capacities originating from the research efforts of the group around Robert Wallerstein in San Francisco. An important criterion for showing the unique contribution of psychoanalytic therapy is the demonstration of changes in the manner of psychological function that cannot be reduced to symptomatic change. As the authors of this chapter emphasise, the prima facie justification for such a measurement approach is clear since psychoanalysis is one of the only current therapeutic interventions that does not aim at the attainment of symptomatic change. The work shows that change beyond the symptomatic may be reliably measured. What value such changes represent in terms of the depressed patient's long-term functioning remains to be established.

The Munich Psychotherapy Study is undoubtedly the most carefully conducted study to date designed to address the vexing question of how, if at all, intensive psychodynamic therapy (psychoanalysis) differs in its effects from psychodynamic psychotherapy. The study is remarkable because it uses a randomised controlled design and because all therapists are highly experienced. The strength of the study which the present contribution highlights is the use of a measure specifically designed to identify changes that may take place in psychoanalysis but are not so characteristic of changes observed following psychoanalytic psychotherapy. The establishment of the validity of such a measure is of paramount importance. The Scales of Psychological Capacities were designed to measure structural change. The SPC is a clinician coded measure and thus blindness in the ratings will be very important to demonstrate. The present contribution, however, clearly establishes its relative independence from simpler measures of adaptation (GAF) as well as symptom severity. The inter-rater reliability of the measure is good, and some progress is reported

towards establishing construct validity and convergent validity by showing depression-specific abnormality in the measure in a group of depressed patients and correlations with self-report measures of interpersonal functioning that might be predicted for this group. The report speaks volumes to the careful way in which this extremely important study is being conducted.

The fluidity of the current psychoanalytic research knowledge base is well illustrated by the contrary position adopted by Ralph Sandell's Karolinska Institute (STOPP) project. The Stockholm study is by far the largest of the prospective investigations of psychoanalysis reported to date. The advantage of intensive compared to non-intensive treatment was interestingly clearest in the symptomatic domain. Thus, the jury is still out on the empirical question of whether measuring change beyond symptoms (as sought by many including the Munich group) is indeed to be the touchstone of psychoanalytic therapy research, or whether the most effective demonstrations are the simplest and a focus on symptom measurement is sufficient. But the STOPP study also highlights the third point of our psychoanalytic trident of effectiveness research program outlined above. The study shows that the superiority of psychoanalysis over psychotherapy, in the long term, is clearest when unmodified 'classical' psychoanalytic ideas govern psychotherapeutic interventions. This approach might be almost ineffective when administered non-intensively. If the ideology of the therapist is broader, the superiority of psychoanalysis over psychotherapy is less marked. Whilst these findings concern ideology rather than actual technique, they do highlight the need to further evolve applications of psychoanalysis, particularly since many of those trained to practice psychoanalysis now often only practice psychotherapy after qualification.

The first contribution from London, from St Anne's Hospital, is in the same spirit of broadening ideology and applications. This is a randomised controlled trial that served to advance the understanding and treatment of treatment of borderline personality disorder (defined psychiatrically) as well as demonstrate remarkably successful long-term outcomes. An important point to note about this study is the extensive use of nurse practitioners in this psychotherapeutic Day Hospital intervention. Some years ago, these therapists might have been treated with condescension by psychoanalytically trained psychiatric colleagues. Now they represent almost the only controlled evaluation of a psychoanalytically oriented intervention with this group. The immense pragmatic importance of Anthony Bateman's work is due to his translation of basic psychoanalytic principles to enable practitioners who would be considered untrained by

traditional standards to administer a systematic and powerful intervention. If psychoanalysis is going to survive in a statutory service that is distributed on a principle of equity, such an ‘enabling’ approach must inevitably be at the core of technical adaptation.

The second contribution from London describes research on an inpatient service for a relatively dangerous group of patients, which Kingsley Norton and Bridget Dolan have directed with great effectiveness. This unit tackles some of the most difficult personality disordered patients: young men whose psychopathology is combined with criminal tendency and significant dangerousness. The report is a case study in itself, showing how retrospective studies of outcome can develop into prospective studies as definitions of improvement are refined. An important contribution of this study is the cost-effectiveness data that the Henderson hospital was able to compile, which undoubtedly contributed to the Henderson model being adopted in a number of other UK settings. Such data is not often available for psychoanalytic psychotherapy evaluations (Gabbard, Lazar, Hornberger, & Spiegel, 1997). The controlled study reported in the chapter suggests important benefits from the programme in the domains of both symptomatology and mood, and underscores the value of the service in that mood variables such as irritability appear not to improve spontaneously at all, but in fact show a slight tendency to worsen in the absence of the therapeutic community provided by the Henderson.

This is a pioneering volume of work in progress. It is important and exciting work by talented pioneers who have responded effectively to an intellectual as well as a professional call. It is clear from the variety of findings reported in this stimulating volume that many of the traditional ideas concerning psychoanalytic psychotherapy will need to be revised. This does not signal the demise of the psychoanalytic approach but rather indicates the great potential for further development of its knowledge base. In the past, psychoanalysis as a theory has not benefited markedly from the rapid virtuous cycle of theoretical development leading to increasingly refined observation and data collection which in turn produces findings that raise theoretical questions that in turn lead to further scientific hypotheses of increased specificity and so on. This book is a signal that this process has finally begun. Under the benevolent nurturing editorship of Phil Richardson and his two co-editors, an excellent sampler has been provided for those who wish to engage in the excitement of systematic data gathering that remains the hope of a future for psychoanalysis.

Peter Fonagy

References

Chiesa, M., & Fonagy, P. (1999). From the efficacy to the effectiveness model in psychotherapy research: The APP multi-centre project. Psychoanalytic Psychotherapy, 13, 259-272.

Cierpka, M., Bucheim, P., Freyberger, H. J., Hoffman, S. O., Janssen, P., Muhs, A., Rudolf, G., Rüger, U., Schneider, W., & Schüssler. (1995). Die erste Version einer Operationalisierten Psychodynamischen Diagnostik (OPD-1). Psychotherapeutics, 40, 69-78.

Cooper, A. M. (1996). Psychoanalysis in the 21st century: Unity in plurality. Paper presented at the 50th anniversary of the Los Angeles Psychoanalytic Society and Institute, Los Angeles, September.

Cornwell, J. (1996). The Power to Harm: Mind, Medicine, and Murder on Trial. London: Viking.

Department of Health. (1995). Strategic Report on the Psychotherapies. London: HMSO.

Detre, T., & McDonald, M. C. (1997). Managed care and the future of psychiatry. Archives of General Psychiatry, 54, 201-204.

Endicott, J., Spitzer, R., Heiss, J., & Cohen, J. (1976). The global assessment scale. A procedure for measuring overall severity of psychiatric disturbance. Arch. Gen. Psychiatry, 33, 766-771.

Faulkner, A. (2000). Strategies for living: A report of user-led research into people's strategies for living with mental distress. London: Mental Health Foundation.

Fonagy, P. (in press). The development of psychopathology from infancy to adulthood: the mysterious unfolding of disturbance in time. Infant Mental Health Journal.

Gabbard, G. O., Gunderson, J. G., & Fonagy, P. (in press). The place of psychoanalytic treatments within psychiatry. Archives of General Psychiatry.

Gabbard, G. O., Lazar, S. G., Hornberger, J., & Spiegel, D. (1997). The economic impact of psychotherapy: A review. American Journal of Psychiatry, 154, 147-155.

Giddens, A. (1999). Runaway world: How Globalisation Is Reshaping Our Lives. London: Profile Books.

Health Council of the Netherlands. (2001). The efficiency of long-term psychotherapy (2001/08E). The Hague: Health Council of the Netherlands.

Hertzman, C., & Wiens, M. (1996). Child development and long-term outcomes: a population health perspective and summary of successful interventions. Social Science and Medicine, 43, 1083-1095.

Knapp, M. R. J. (1997). Economic evaluations and interventions for children and adolescents with mental health problems. Journal of Child Psychology and Psychiatry, 38(1), 3-25.
Markowitz, J. C., & Street, L. L. (1999, October 15, 1999). NIMH propels psychotherapy on a new course. Psychiatric News.
Mayes, N., & Pope, C. (2000). Qualitative research in health care. Assessing quality and qualitative research. British Medical Journal, 320, 50-52.
Menninger, K., Mayman, M., & Pruyser, P. (1963). The vital balance: The life process in mental health and illness. New York: Viking Press.
Michels, R. (1994). Psychoanalysis enters its second century. Annual of Psychoanalysis, 22, 37-45.
O'Connor, T., Bredenkamp, D., Rutter, M., & team, a. t. E. a. R. a. s. (1999). Attachment disturbances and disorders in children exposed to early severe deprivation. Infant mental Health Journal, 20, 10-29.
O'Connor, T., Rutter, M., & team, a. t. E. a. R. a. s. (in press). Attachment disorder behavior following early severe deprivation: Extension and longitudinal follow-up. Journal of the American Academy of Child and Adolescent Psychiatry.
Rose, D. (2001). User's voices: The perspective of mental health service users on community and hospital care. London: Sainsbury Centre for Mental Health.

Rutter, M. (2000). Psychosocial influences: Critiques, findings and research needs. Development and Psychopathology, 12, 375-405.
Sackett, D. L., Rosenberg, W. M., Gray, J. A. M., Haynes, R. B., & Richardson, W. S. (1996). Evidence based medicine: What it is and what it isn't. British Medical Journal, 312, 71-72.
Shaffer, D., Gould, M. S., Brasic, J., Ambrosini, P., Fisher, P., Bird, H., & Aluwahlia, S. (1983). A children's global assessment scale (CGAS). Archives of General Psychiatry, 40, 1228-1231.
Shea, M. T., Elkin, I., Imber, S. D., Sotsky, S. M., Watkins, J. T., Collins, J. F., Pilkonis, P. A., Beckham, E., Glass, D. R., Dolan, R. T., & Parloff, M. B. (1992). Course of depressive symptoms over follow-up: Findings from the NIMH treatment of depression collaborative research programme. Archives of General Psychiatry, 49, 782-787.
Solms, M. (2000). Dreaming and REM sleeping are ocntrolled by different brain mechanisms. Behavior and Brain Sciences, 23, 843-850; 904-1121.
Solms, M., & Nersessian, E. (1999). Freud's theory of affect. Neuro-psychoanalysis, 1, 5-14.

Weisz, J. R., Hawley, K. M., Pilkonis, P. A., Woody, S. R., & Follette, W. C. (2000). Stressing the (other) three Rs in the search for empirically supported treatments: review procedures, research quality, relevance to practice and the public interest. Clinical Psychology: Science and Practice, 7(3), 243-258.
Weisz, J. R., & Jensen, P. S. (1999). Efficacy and effectiveness of child and adolescent psychotherapy and pharmacotherapy. Mental Health Services Research, 1, 125-157.
Wells, K. B. (1999). Treatment research at the crossroads: The scientific interface of clinical trials and effectiveness research. The American Journal of Psychiatry, 156, 5-10.
Westen, D., & Morrison, K. (in preparation). How empirically valid are empirically validated therapies? A critical appraisal. Unpublished manuscript.

Wing, J., Curtis, R., & Beevor, A. (1996). HONOS: Health of the Nation Outcome Scales - a report on research and development (July 1993 - December 1995). London: College Research Unit, Royal College of Psychiatrists.
Wing, J. K., Lelliott, P., & Beevor, A. S. (2000). Progress on HoNOS. British Journal of Psychiatry, 176, 392-393.

The Influence of psychodynamic psychotherapy on personality traits of Anorexia nervosa and Bulimia nervosa in-patients – results of the German multi-centre eating disorder study (TR-EAT)

Jörn von Wietersheim[1], Peter Malewski[2], Burkard Jäger[2], Werner Köpp[3], Ines Gitzinger[1], Peter Köhler[4], Ralf Grabhorn[5] , Horst Kächele[1] & TR-EAT

[1] Department of Psychotherapy and Psychosomatic Medicine, University of Ulm, Germany
[2] Department of Psychosomatic Medicine, University Hospital Hannover, Germany
[3] Department of Psychosomatic Medicine and Psychotherapy, University Hospital Steglitz, Free
University Berlin, Germany
[4] Department of Psychosomatic Medicine, Medical University Lübeck, Germany
[5] Department of Psychosomatic Medicine, Hospital of the Johann-Wolfgang-Goethe-University Frankfurt,
Germany

Correspondence address: PD Dr. Jörn von Wietersheim, Sektion Psychosomatische Medizin, Universitätsklinikum Ulm 89070 Ulm
Tel: 0731 / 50-24873

Fax: 0731 / 50-24535
e-mail: joern.vonwietersheim@medizin.uni-ulm.de

Funded by The Ministry of Education and Research (BmBF, FKZ01EN9410)

The Influence of psychodynamic psychotherapy on personality traits of Anorexia nervosa and Bulimia nervosa in-patients – results of the German multi--centre eating disorder study (TR-EAT)

Objectives: What are the effects of a psychodynamic in-patient treatment of patients with anorexia nervosa and bulimia nervosa on personality data? In how many patients can a successful change in the personality domain be observed?

Methods: 732 patients were assessed at the beginning and the end of an in-patient treatment as well as 2_ years after this treatment. Data were collected by means of the personality inventories „Freiburger Persönlichkeitsinventar (FPI-R)“ and „Narzissmusinventar“. A definition of success was constructed using the clinical significance concept. This is related to the scales life satisfaction, inhibition (FPI-R), powerless self and negative body self (Narzissmusinventar).

Results: The results of both questionnaires reflect clinically well-known psychopathological characteristics of eating disorder patients. During in-patient treatment and also afterwards, there were improvements in personality data, but not to the level obtained from a healthy control group. One third of the patients showed remarkable improvements in the outcome measure. The success in the personality field is positively correlated to the success in the eating disorder symptoms. This success could not be predicted by initial data or by therapy data.

Conclusions: There are marked impairments of the patients in the personality data. Changes in this domain take time and are rather small.

Introduction

The eating disorders anorexia nervosa, bulimia nervosa and the corresponding double diagnoses in the diagnosis schemes DSM III-R, DSM IV as well as ICD-10 are predominantly defined by symptoms. Additionally, numerous publications (e.g. Feiereis, 1989, Senf, 1989, Janssen et al., 1997, Herzog et al., 1995) but also clinical experience have demonstrated that these patients show significant instabilities and disorders with respect to personality traits. Many patients have been reported to suffer from depressive moods, anxieties, low self-esteem, tendencies to social isolation, and, particularly for anorectic patients, compulsive behaviour. As a result, treatments (psychodynamic psychotherapies in particular) seek to tackle these personality characteristics. Treatment of current conflicts as well as strengthening of the resources of individual patients are important goals, besides reducing the symptoms. Consequently, we designed a study, which assessed success not only in terms of symptoms, but also in terms of personality and interpersonal relations (Kächele, 1999, Kächele 2000). This report focuses on the measurement of success with respect to personality traits. Personality questionnaires are commonly used to record personality traits and characteristics. High scores on depression, social isolation, enhanced psychopathology, inhibition, powerless self-esteem and reduced life satisfaction have been reported in previous studies on eating disorders, using different questionnaires (Böhle et al. 1991, Schork et al. 1994, Jäger et al. 1996, Hurt et al. 1997, Dancyger et al. 1997, Thiel et al. 1999). Anorectic patients (purging type) show higher psychopathology than restrictive type patients (Hurt et al., 1997). Dancyger et al. (1994) describe marked changes of the psychopathology during an in-patient treatment and fewer changes during the follow-up period of 10 years.

This multi- centre eating disorder study (Project TR-EAT) emerged from a combined research effort of various specialised hospitals and university hospitals. The goal of the study was to investigate the courses of anorexia nervosa and bulimia nervosa patients. The focal point of the research was to find out whether or not different periods of in-patient treatment affect the therapy outcome after 2 _ years. Further interest was related to the prognosis of these diseases and the possibility of prediction of success at the follow-up evaluation. Numerous symptomatic, but also personality-diagnostic and social variables have been recorded by questionnaire and interview at the beginning and at the end of an in-patient treatment, and also at the follow-up examination 2_ years after beginning of the treatment. 45 clinics participated in this study and 1247 female patients were assessed in total. As a main result it emerged that 36 % of the anorexia-

patients and likewise 36 % of the bulimia-patients were largely free of symptoms at the follow-up assessment. The duration (between 6 weeks and up to more than 3 month) and intensity (number of psychotherapy sessions) of the in-patient treatment varied greatly. However, these variations did not influence the outcome after 2 _ years. More then 80 % of the patients continued psychotherapy after the in-patient treatment, mostly as outpatients, but sometimes as in-patients in another hospital.

However, a distinction must be made between research into personality traits and studies dealing with personality disorders of patients with eating disorders. Personality disorders are well defined (e.g. according to DSM-IV or ICD-10) diagnoses, which can be drawn in addition to an existing eating disorder and describe a heavy psychopathology. Accordingly, the co-morbidity of anorexia and personality disorders has been estimated to be up to approximately 50 % (Rosenvinge and Mouland, 1990). Estimations with bulimic patients with additional personality disorders range at ca. 20 % (Herzog et al., 1995). For the assessment of treatment success the use of a personality disorder comorbidity is likely to be less sensitive than the standardized assessment of personality traits.

2. Method

This work is a partial evaluation carried out in the context of the multi- centre eating disorder study (Kächele *et al.*, 2000, Kächele *et al.*, 1999). Only the methodological aspects relevant for this evaluation are considered here.

The instruments used for personality data collection were the *Freiburger Persönlichkeitsinventar (FPI-R)* (Fahrenberg *et al.*, 1989) and the *Narzissmusinventar* (*Narcism-inventory*, Denecke and Hilgenstock, 1989). The FPI-R was set as the German standard assessment instrument for personality data. Since problems of the narcissistic regulatory system are commonly associated with eating disorder patients, the *narcissism inventory* has also been used. For economical reasons, however, a third of the scales have been discarded, as no deviation from a ‘healthy’ control group could be found in previous studies. The internal consistency (Cronbach’s-Alpha) of the narcissism inventory scales lies between 0.71 and 0.94. Overlap with the FPI-R is only marginal.

Data on personality were recorded at the beginning and the end of an in-patient treatment, which on average lasted approximately 3 months, as well as at the

time of the follow-up evaluation 2_ years after admission into the in-patient treatment.

Four levels of symptomatic success were defined according to the DSM III-R-diagnostic criteria. The components used in each case were the main symptom, i.e. for anorexia nervosa being underweight (15 % below the expected weight), for bulimia binging (at least twice per week), as well as other symptoms such as body image distortion, fear of fatness, and weight reduction strategies. Success assessment was scaled as follows: 0 refers to the assessment „all relevant symptoms are pathological“, 1 „one symptom in the healthy range“, 2 „the main symptom (weight or binges, respectively) and one additional symptom in the healthy range“, 3 „all symptoms in the healthy range“.

It is very problematic and doubtful to define a single criterion for success for the personality field. Based on experience with these instruments, a criterion has been developed, which contained four scales that were considered necessary. Those were from the FPI-R scales for life satisfaction and self-consciousness and from the narcissism inventory the scales for powerless self and negative body-self. These scales were selected as they reflect important aspects of specific problems of patients with eating disorders on one hand, and on the other hand cover different content areas. A further consideration regarding the definition of success was that the main goal of therapy should be to improve pathological (deviating from the normal range) values in the relevant scales, and to reach norm-values if possible. Hence, if a value falls outside the normal range before but inside after treatment, this should be considered a success (corresponding to one point in the success criterion). If a value improves clinically significantly, but still lies outside the normal range (e.g. improvement of Stanine values from 1 to 3), it should be considered a partial success (half point). Here, a definition of a clinically significant change, which is related to the reliability of an instrument (questionnaire), has been used (Jacobson and Truax, 1991). Consequently, a criterion to measure success has been developed, which considered both, absolute changes (outside before treatment and inside after) and relative changes (strongly pathological before treatment, clinically significantly less pathological after). If a value has been in the normal range right from the start, the therapy was judged successful concerning the respective scale (this could lead to an over estimation of the effectiveness of therapy, but this rarely occurred). The success points were added up and a metric standard for success was constructed with them. The value 6 equalled maximum success reflecting

maximum changes, the value 0 corresponds to no changes in personality values, and negative values (to a maximum of –2) indicated deterioration.

All patients were admitted to one of 45 psychodynamically orientated psychotherapeutic hospitals in Germany between September 1993 and October 1995. The 2-_ years follow-up assessment was conducted by the end of 1998. Inclusion criteria were DSM III-R-diagnosis for anorexia nervosa, Bulimia nervosa and double-diagnosis (both anorexia- and bulimia criteria fulfilled) as well as an age of ≥18 years. 1247 mostly female patients participated in this study, however, the sample size decreased owing to data inconsistencies to 1171 participants.

The following description of the samples is for the 732 patients who participated in the follow up. 712 patients were female, 20 were male; 405 were bulimic, 229 anorectic, and 98 cases were double-diagnoses. The average age of anorectic patients was 24.9 (SD = 6,0) years, for bulimic patients 25,9 (SD = 6,3), and for patients with double-diagnosis 25.4 years (SD = 5,7). The average duration of illness for anorectic patients was 5.8 years (SD = 5,4), for bulimic patients 8.2 years (SD = 6,2), and for patients with double-diagnosis 6,7 years (SD = 5,5). The average duration of in-patient treatment was 12.2 weeks (SD = 7,7).

As yet, there are no normative data available for the narcissism inventory. Thus, data from a control group, consisting of 120 female medical students, have been used[1].

3. Results

3.1 Freiburger personality inventory

Table 1 depicts the FPI-R-stanine values for the anorexia- and bulimia group. Many scales show changes in the means, which numerically (and probably also clinically) are rather small. Particularly obvious are very low life satisfaction, high inhibition, high arousal, and high emotionality (Neuroticism). Changes of the means in the desired direction became apparent in the previously defined scales for success measurement. Also, many of the patients show clinically deviating personality characteristics (stanine values < 4 and > 6 respectively). Changes take place in the period from admission to discharge as well as from discharge to the date of follow-up evaluation. Patients with double-diagnosis, which, owing to lack of space, are not included in Table 1, displayed particularly

[1] These data are a courtesy of PD Dr. K. Engel, Dortmund.

low values for “life-satisfaction” but also particularly high values for “strain” and ”emotionality“. These patients seem to have a higher psychopathology.

Table 1: FPI-scales (Stanine) for „admission“, „release“ and „follow-up“.

FPI	Anorexia			Bulimia		
	Admission	**Discharge**	**Follow-up**	**Admission**	**Discharge**	**Follow-up**
Life satisfaction	2.7 ±1.4	3.1 ±1.6	3.5 ±1.8	2.7 ±1.2	3.1 ±1.4	3.6 ±1.8
Social orientation	5.9 ±1.8	5.5 ±1.7	5.1 ±1.8	5.6 ±1.8	5.1 ±1.9	5 ±1.8
Efficiency orientation	4.8 ±1.9	5.2 ±1.9	5.1 ±1.9	4.5 ±1.9	4.9 ±1.9	4.8 ±2
Inhibition	6.8 ±1.9	6.4 ±1.8	6.1 ±2	6.4 ±2	6 ±2	5.9 ±2.1
Arousal	6.4 ±1.7	6.3 ±1.8	6.2 ±1.8	6.4 ±1.8	6.3 ±1.8	6.2 ±1.9
Aggressiveness	4.6 ±1.9	4.8 ±1.9	4.9 ±1.7	5.2 ±1.9	5.4 ±2	5.4 ±1.7
Strain	6.1 ±1.5	5.8 ±1.6	5.7 ±1.7	6.1 ±1.6	5.8 ±1.7	5.7 ±1.8
Physical complains	6.8 ±1.9	6.1 ±2	6.2 ±2.1	6.7 ±1.7	6.1 ±1.8	5.7 ±2
Health concerns	4.3 ±2.1	4.3 ±2.1	4 ±2.1	3.7 ±1.8	3.9 ±2	4 ±2
Openness	5.3 ±1.8	5.4 ±1.8	5.3 ±1.8	6 ±1.8	6 ±1.9	6 ±1.7
Extraversion	3.4 ±1.9	4 ±1.9	3.9 ±1.9	4.1 ±2.1	4.6 ±2.1	4.4 ±2
Emotionality	7.1 ±1.6	6.8 ±1.9	6.6 ±2	7.3 ±1.4	6.9 ±1.7	6.6 ±2

The means and standard deviation are shown (Anorexia N=229, Bulimia N=405).

3.2 Narcissism inventory

Table 2 shows the scales of the narcissism inventory. Here too, changes in the mean values became apparent, even more than with FPI-R.

Table 2: Scales of the narcism inventory for "admission", "release" and „follow-up".

Narcism inventory	Anorexia			Bulimia		
	Admission	**Discharge**	**Follow-up**	**Admission**	**Discharge**	**Follow-up**
Powerless self	32.3 ±9.7	27.2 ±10.1	27.1 ±11.2	33.1 ±9.5	27.3 ±10	26.1 ±10.9
Loss of Affect-Impulse-Control	30.7 ±8.5	28.6 ±8.9	27.9 ±9.1	33.3 ±8.8	30.2 ±8.6	29.2 ±9.1
De-realisation/ De-personalisation	29.9 ±10.7	25.3 ±10.6	24.7 ±11.7	29.8 ±10.5	25 ±10.3	24 ±11.1
Basic hope-potential	30.2 ±9.3	32.1 ±9.6	32.3 ±9.5	28.9 ±9.4	32 ±9.6	33.8 ±9.7
Diminished self-image	35.1 ±8.7	31.9 ±9	32.3 ±9.3	34.9 ±8.8	31.2 ±8.9	30.3 ±9.3
Negative body-self	30 ±11.1	24.5 ±10.9	25.2 ±12.5	31.2 ±11.4	24.7 ±11.4	24 ±12.2
Social isolation	30.8 ±8.6	28.5 ±8.6	28.3 ±8.7	29.4 ±8.6	27.5 ±8.2	27.1 ±9
Greatness self	24.3 ±7	26.1 ±6.9	26.7 ±6.8	26.4 ±7.4	28.2 ±7	28.8 ±7.2
Narcistic rage	26.8 ±8.1	26.6 ±7.7	26.4 ±7.6	28.1 ±8.1	27.8 ±7.6	27.3 ±7.8
Object devaluation	26.9 ±7.1	25.4 ±7.3	26.7 ±7.7	28.7 ±7.2	26.7 ±7.5	27 ±7.6
Symbiontic self-protection	38.7 ±5.7	36 ±6.3	37.4 ±6.4	37.6 ±6.5	35.4 ±6.6	36.7 ±6.5
Hypochondric anxiety bonds	26.1 ±9.4	23 ±9.3	22.7 ±9.6	25.7 ±9	22.5 ±8.5	23.1 ±9.1
Narcistic Illness gain	24.6 ±9.3	22.6 ±9.3	21.4 ±10.3	21.4 ±8.7	19 ±8.1	19.4 ±8.9

The means and standard deviation are shown (Anorexia N=229, Bulimia N=405).

Significant changes can be seen between admission and discharge on the scales 6 "negative body-self", 1 "powerless self", 3 "de-realisation/ depersonalisation", and 5 "diminished self-image". Those effects became apparent in all three diagnostic groups. Accordingly, the patients know to accept their bodies slightly better after in-patient treatment, they feel less powerless and inferior. Over all, changes rather seem to take place between in-patient admission and discharge than between discharge and the follow-up assessment, although most patients experienced outpatient psychotherapy after in-patient treatment. Clearly, distinct effects of intense in-patient psychotherapy can be seen from these questionnaires.

A comparison between the groups of patients with eating disorders and the control group (120 female medical students) revealed significant deviations on many scales. Particularly significant differences emerged for the scales powerless

self, de-realisation, diminished self-image, negative body- image, and social isolation. Known clinical signs and symptoms are hence well reflected in the questionnaire data for these patients.

3.3 Calculation and distribution of the success criterion

Figure 1 shows the number of patients that deviated in the four target criteria scales from the normal range at the point of admission and follow-up assessment. The largest group deviates from the normal range on all four target scales. In contrast 4 % of patients fell within the normal range on these scales (despite an apparent eating disorder). Some change is visible from the time of admission to follow-up assessment. A few patients, who previously were outside the defined normal range, now fall within. In all, only about 20 % of patients experience this degree of change, with the majority remaining within the pathological range on the scales.

Fig. 1: Number of patients and personality scales outside the norm

Note: Recorded are the FPI scales for life satisfaction and arousal as well as from the narcism inventory the scales powerless self and negative body-self.

The success criterion used in this study has been developed based on the clinically significant change criterion of Jacobsen and Truax, (1991). It is composed of absolute values (whether or not a patient is within the normal range of a respective scale) and the values for changes (whether or not a patient changed clinically significantly since admission). Figure 2 shows the distribution of the success criterion for the groups of anorexia and bulimia nervosa patients. A relatively large group of patients (13 %) can be seen who do not show any evidence of change. A further 13 % of the patients showed evidence of deterioration during the course of treatment, the majority of the patients however showed improvements, which show a relatively wide deviation. Approximately a third of the patients show distinct improvement (value of success ≥ 3). Comparing the three diagnosis groups it becomes apparent that anorexia and bulimia patients show similar success ratios, whereas patients with double diagnosis show significantly less evidence of success.

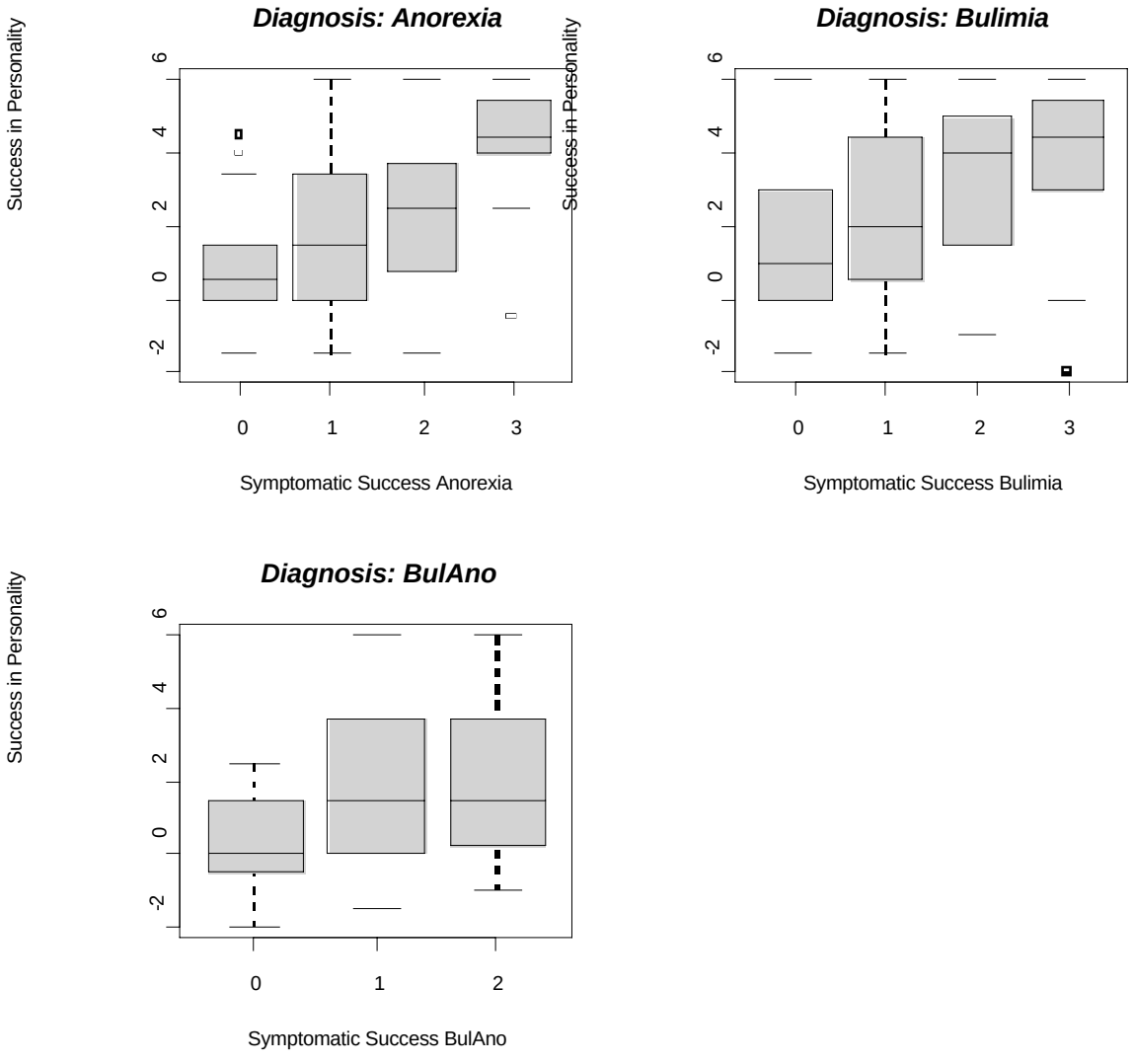
Fig. 2: Distribution of percentages for the success criterion "personality" for each diagnosis group.

Note: The criterion is the sum of the scales, which at the moment of follow up were within the norm (coded [0,1]) and the number of scales that changes clinical significant since admission (coded [0, _]).

3.4 Relations between the criteria for improvement on personality and symptommeasures

These relations are shown in figure 3 (the symptomatic criterion in four levels, the personality criterion scaled to the interval). Good agreement can be seen from these illustrations. They are however not really linear, as the respective correlations (Spearman's Rho) for anorexia was 0.52, for bulimia 0.35 and for double diagnoses 0.27. According to that, change in personality corresponds with changes in symptoms. The direction of causality, if any, in this relationship remains unclear.

Fig. 3: Relation between symptomatic success and success in personality scales



Note: Symptomatic success: 0 = all relevant symptoms are outside the norm
1 = 1 symptom inside the norm,
2 = main symptom (weight or binges) and another symptom inside the norm
3 = all relevant symptoms inside the norm
No patient of the group with double diagnosis (BulAno) reached the value 3 for symptomatic success.

Differential Treatment Outcome of Inpatient Psychodynamic Group Work

3.5 Prediction

To what extent is prediction of a positive outcome possible on the personality measures? For this, two models have been modified and successively tested. The first model tests the association of relevant patient traits with the degree of patient improvement. The second model tested, whether or not the therapy parameters (duration of in-patient treatment, number of therapy sessions) influenced the degree of improvement. To compensate for model violations, the response-variable has been transformed according to the BoxCox-transformation (lambda=. 24).

Two combined measures, which characterised the symptomatic state at admission, have been applied as potentialpredictors. Further predictors include the extent of pre-treatment, the duration of illness, the patient's weight (BMI), their desired weight, the overall SCL-90-value, as well as the sum of the three EDI scales. The calculated first model, however, only clarifies 5 % of the variance (adjusted). For the second model, the amount of therapy was considered, as the total number (duration of in-patient treatment in days) and the weekly "dose" in hours. Likewise within this expanded model, only 5 % of the variance could be accounted for. Accordingly, prediction before treatment as to later therapeutic success in the personality domain was not possible, at least with the variables employed in the present study.

4. Discussion

Both, the Freiburger Personality Inventory as well as the Narcissism Inventory reflect clinically known psychopathological signs and symptoms of patients with eating disorders. It has to be emphasised however, that regardless of their ascertained diagnoses of eating disorder, not all patients deviated in the questionnaire values from the normal range. There were a small number of eating disorder patients with normal questionnaire values. It is possible that these patients were dissimulating. But it is also conceivable, that they developed anorectic or bulimic reactions due to an acute crisis, with the personality in fact remaining relatively unaffected. The score distributions are relatively broad; i.e. the patients are very different. Changes, which are reflected in the questionnaire values, occurred during in-patient treatment but also during subsequent outpatient treatment.

Nevertheless the values of most patients did not normalise during the observation period. They felt slightly better but the majority did not reach the values of a norm-group (FPI-R) or a healthy control group in the narcissism inventory, respectively. This can be seen to be paralleled by the successful outcomes wheresymptoms were concerned: most patients showed marked eating disorder symptoms after 2 _ years of observation, only about 36 % of which were considered clinically healthy (Kächele 2000).

These slow and limited changes in the questionnaire values of the FPI and narcissism inventory match theoretical considerations as to personality and eating disorders well. Personality is considered an enduring construct, which evolves and changes only very slowly. In addition clinical experience from work with patients with eating disorders corresponds with these results: changes only happen gradually. It is clear that change on the narcissism inventory appear more frequently during the three months (on average) of in-patient treatment, whereas fewer changes occur in the follow-up period, when most patients received outpatient therapy. These findings support and justify such intensive psychotherapeutical procedures as in-patient treatment. Dancyger et al. (1997) reported very similar results. In the latter case, anorexic patients were studied by means of the MMPI at hospital admission, discharge and follow-up evaluation.

Direct comparisons between change (successful outcomes) in the fields of personality and symptoms show moderate agreement. Symptomatic improvement went along with improvements in the domain of personality; the correlation coefficients (Spearman’s Rho), however, were only of moderate strength. The

question of whether changes in the fields of personality and apparent symptoms coincide or take place in succession, and if change in one field (e.g. personality) is a requirement for change in the other (e.g. symptoms), remains unclear and will be the subject of future investigation.

Successful prediction of improvement in the personality domain could not be achieved. Change in personality could also not be explained by patient variables such as degree of apparent symptoms, duration of illness, number of previous treatments etc., or therapy variables such as duration and intensity of in-patient therapy. The variance accounted for in each case was only around. 5 %. These results correspond to those reported by Kächele (2000), which are related to success regarding apparent symptoms. Also from this work, no safe method for prediction emerged.

Consequently, it remains unclear with what the changes of in-patients with eating disorders are connected. This problem should be the subject of future investigation, which should include a discussion on the relatively short duration of follow-up. It is conceivable that many patients after 2 _ years are still in the process of symptomatic and personality related development, and recognisable effects only become apparent at a later stage of the course of their illness.

References:

Böhle A., Wietersheim J. von, Wilke E., Feiereis H. (1991): Soziale Integration bei Anorexia nervosa und Bulimie. Zsch Psychosom Med Psychoanal 37, 282-291.

Dancyger I. F., Sunday S. R., Eckert E. D, Halmi K.A. (1997): A comparative analysis of Minnesota Multiphasic Personality Inventory profiles of anorexia nervosa at hospital admission, discharge, and 10-year follow up. Compr Psychiatr 38, 185-191.

Deneke F. W., Hilgenstock B. (1989): Das Narzissmusinventar. Bern: Huber.

Fahrenberg J., Hampel R., Selg H. (1989): Das Freiburger Persönlichkeitsinventar FPI. 5 Aufl. Göttingen: Hogrefe.

Fahy T. A., Eisler I., Russell G. F. (1993): Personality disorder and treatment response in bulimia nervosa. Br J Psychiatry 162, 765-770.

Feiereis H. (1989): Diagnostik und Therapie der Magersucht und Bulimie. München: Marseille.

Herzog T., Stiewe M., Sandholz A., Hartmann A. (1995): Borderline-Syndrom und Essstörungen. Literaturübersicht und Interview-Studie an 172 konsekutiven Inanspruchnahmepatientinnen der Freiburger Essstörungsambulanz. Psychother Psychosom Med Psychol 45, 97-108.

Hurt S. W., Brun-Ebérentz A., Commerford M. C., Samuel-Lajeuness B., Halmi K. A. (1997) A comparison of psychopathology in eating disorder patients from France and the United States. Int J Eat Disord 22, 153-158.

Jacobson N. S., Truax P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy. J Consult Clin Psychol 59, 12-19.

Jäger, B., Liedtke R., Künsebeck H.-W., Lempa W., Kersting A., Seide L. (1996): Psychotherapy and bulimia nervosa: evaluation and long-term follow-up of two conflict-orientated treatment conditions. Acta Psychiatr Scand 93, 268-278.

Janssen P. L., Senf W., Meermann R. (1997): Klinik der Essstörungen: Magersucht und Bulimie. Stuttgart: Fischer.

Kächele, H. und MZ-ESS (1999): Eine multizentrische Studie zu Aufwand und Erfolg bei psychodynamischer Therapie von Essstörugen. Studiendesign und erste Ergebniss. Psychother Psychosom Med Psychol 49, 100-108.

Kächele, H. und MZ-ESS (2000): Aufwand und Ergebnis stationärer psychodynamischer Psychotherapie von Essstörungen in Deutschland. Projektbericht. Stuttgart: Forschungsstelle für Psychotherapie. Submitted to Psychotherapy Research.

Rosenvinge, J. H., Mouland, S. O. (1990): Outcome and prognosis of anorexia nervosa. A retrospective study of 41 subjects. Br J Psychiatry 156, 92-97.

Senf W. (1989): Psychoanalytische Betrachtungen zur Bulimie. In Kämmerer A., Klingenspor B. (Hrsg): Bulimie- Zum Verständnis einer geschlechtsspezifischen Erkrankung. Stuttgart, Kohlhammer, S. 88-103

Schork E. J., Eckert, E. D., Halmi K. A. (1994): The relationship between psychopathology, eating disorder diagnosis, and clinical outcome at 10-year follow-up in anorexia nervosa. Compr Psychiatr 35, 113-123.

Thiel A., Züger M., Jacoby G.E., Schüßler G. (1999): Welche Bedeutung haben narzißtische Selbst-Störungen für die Prognose der Anorexia und Bulimia nervosa? Eine prospektive Katamnese über 30 Monate. Zsch psychosom ed Psychoanal 45, 57-76.

Differential Treatment Outcome of Inpatient Psychodynamic Group Work

Bernhard Strauss, Silke Schmidt
Institute of Medical Psychology, University Hospital,
Friedrich-Schiller-University Jena, Germany

Correspondence concerning this article should be sent to: Prof. Dr. Bernhard Strauss, Institute of Medical Psychology, University Hospital Friedrich-Schiller-University, Jena Stoystrasse 3, D 07740 Jena, Germany
bernhard.strauss@med.uni-jena.de

Background: The German Working Group on Inpatient Group Psychotherapy

For historical reasons psychotherapy in Germany has a very specific tradition (cf. Strauss & Kaechele, 1998). Part of this tradition is the high importance of inpatient psychotherapy. Genuine psychotherapeutic hospitals had already been founded in Germany before World War II (e.g., by Georg Simmel in Berlin and

Georg Groddeck in Baden-Baden). After the World War a whole series of psychotherapeutic hospitals were founded as well as several hospitals for psychotherapeutic and psychosomatic rehabilitation. This has contributed to the fact that Germany has still well over 10,000 hospital beds for psychotherapy alone, apart from those within psychiatry. Accordingly, inpatient psychotherapy has become a very important part of the psychotherapy delivery system in this country.

Compared to publications dealing with the development and differentiation of treatment concepts within the inpatient field, empirical studies remained rare until the eighties (Strauss, 1992). With the foundation of the Mainz Workshop on Research in Inpatient Psychotherapy (1988, by M. Bassler and S. O. Hoffmann), empirical research received a new impetus resulting in several initiatives to study the process and outcome of inpatient treatments. Related to the Mainz workshop, several working groups have been established. These working groups tried to focus on specific research questions in the field and to organize cooperation between university hospitals and other psychotherapeutic hospitals. The German Working Group on Inpatient Group Psychotherapy was founded by B. Strauss and J. Eckert (Hamburg University) in the year 1990. The major goal of the research network was to consider the importance of group work within the inpatient setting. In general, compared to individual therapy, there is a considerable lack of studies focussing on the process and outcome of psychodynamic group treatments (cf. Fonagy et al, 1999).

The research group presently consists of colleagues from 10 university hospitals or institutions (Hamburg, Leipzig, Bielefeld, Innsbruck, Kiel, Mainz, Jena, Göttingen, Hannover, Düsseldorf), and 8 general psychotherapeutic or rehabilitation hospitals (Hamburg, Berlin, Grönenbach, Bad Honnef, Geldern, Bad Kreuznach, Frankfurt/Oder, Bad Segeberg). The programme of the group is to organise multi-site studies focussing on very specific questions in the field of inpatient group psychotherapy. In general, it is intended to run these studies in naturalistic settings, with a design that should be easy to integrate into the clinical routine making the raising of specific funds unnecessary.

Since its foundation the working group has completed a variety of studies on the assignment of patients to different subsettings (Eckert, Biermann-Ratjen, Brobeck, et al., 1997), quality assurance in inpatient groups (Strauss, Kriebel, & Mattke, 1998) or the predictive value of attachment styles for the outcome of inpatient groups (e.g. Strauss, Lobo-Drost, & Pilkonis, 1999). One major focus

of the work during the last ten years, which is summarised below concerns the question as to whether the amount and kind of interpersonal problems reported by patients at admission is of any value in differentially predicting the outcome of inpatient psychodynamic group treatment.

The General research question: Do interpersonal problems predict treatment outcome?

Research on inpatient psychodynamic group work has focussed on different questions during the last years (Strauss, 1992), namely the investigation of its global effectiveness (e.g. Strauss & Burgmeier-Lohse, 1994), the subjective importance of single components of the inpatient programme for the patients (e.g. von Rad, Senf, & Bräutigam, 1998), the study of therapeutic factors (e.g. Strauss & Burgmeier-Lohse, 1995) and process research (e.g. Tschuschke, 1993). The working group has primarily focussed on the question of how the patients´ experience of their interpersonal problems – as assessed by the Inventory of Interpersonal Problems (Horowitz, Strauss, & Kordy, 1994) – can predict the outcome of group treatment. This question is based upon the observation that interpersonal relationships, their structure, dynamics and adaptability, play an important role in the course and outcome of psychodynamic psychotherapy. The interpersonal model which has its origins in the psychoanalytical literature (e.g. Horney, 1945) has been shown to be relevant for describing human behaviour in psychotherapy.

In our studies it was assumed that the degree and the quality of interpersonal problems which patients undergoing inpatient group treatment report , should be a differential indicator of treatment outcome. This assumption was based on studies from other fields of psychotherapy showing that interpersonal problems might be significant factors enabling the differentiation of patients with varying degrees of treatment success (e. g. Mohr, Beutler, Engle, et al., 1990; Horowitz, Rosenberg, & Bartholomew, 1993). Before undertaking our reserach , no data were available regarding this question in the inpatient setting. Specifically, we expected that patients who were aware of their interpersonal difficulties should be more susceptible to psychodynamic treatment approaches, whereas those who did not experience interpersonal complaints might be better assigned to other treatment approaches.

Three studies have now been conducted within the working group and these will be summarised in this article. Data related to the most recent study (study 3) have not yet been published, whereas the finding of the other two studies have

also been reported elsewhere (Strauss, Eckert, & Ott, 1993; Davies-Osterkamp, Strauss, & Schmitz, 1996).

Central research method: The Inventory of Interpersonal Problems (IIP)

Based upon the observation that the most frequent complaint of patients upon entering psychotherapy is that they find it difficult to get on with other people, Horowitz and co-workers (1988) compiled the Inventory of Interpersonal Problems (IIP), which has now become a standard instrument in psychotherapy research (Strupp, Horowitz, & Lambert, 1997). The questionnaire originally covered 127 items describing common interpersonal difficulties. Whereas some of these items begin with the statement „It is hard for me ...“ (to do something, e.g. say „no“ to other people), other items indicate that the person does certain things too much (e.g. I fight with other people too much). Horowitz, Rosenberg, Baer et al. (1988) demonstrated that – according to the assumptions of the interpersonal model – the items of the IIP could be arranged along the two orthogonal axes „affection“ (friendliness vs. hostility) and „control“ (dominance vs. submission). Alden, Wiggins, & Pincus (1990) showed that the items fit into a circumplex structure (cf. Figure 1) and constructed eight subscales of eight items each (i.e. a total of 64), each covering an eighth or an octant of the two-dimensional model.

Insert Figure 1 about here

Study 1:

The first study testing the major hypothesis was carried out with the participation of eight hospitals all providing similar forms of psychodynamically oriented group treatment. 470 patients from these hospitals were included in this study in which the results were analysed individually for each institution (for a comprehensive description of the detailed results cf. a special issue of the journal „Gruppenpsychotherapie und Gruppendynamik“ [Group psychotherapy and Group dynamics], Strauss, Eckert, & Ott, 1993).

Of the 470 patients participating in this study, 66% were female. The mean age of the patients was 31 with an age range between 17 and 59 years (SD=11.8) Individual hospitals contributed to the study with patient numbers varying between 24 and 110. Although the concepts of the group treatment were similar, the general treatment programmes differed with respect to the length of the inpatient therapy. The shortest programme lasted between 5 and 6 weeks, the longest between 6 and 7 months. Owing to the fact that the single hospitals used

their own criteria for the assessment of treatment outcome it was hard to integrate the individual findings (cf. Strauss, Eckert, & Hess, 1993). Nevertheless, some general trends could be detected:

Patients suffering mainly from interpersonal distress (as compared to symptomatic distress as measured with the SCL-90-R) had more benefit from the inpatient group treatment in a setting with shorter treatment length (Muhs, 1993), whereas patients predominantly suffering from symptomatic distress had a better outcome in a long-term group treatment (Strauss & Burgmeier-Lohse, 1993).

As far as specific interpersonal problems were concerned, there was a tendency for patients suffering from problems with hostile dominance to have a more negative prognosis than others. This finding was consistent across three of the six hospitals.

Study 2:

Out of the sample of study 1, a subsample of 194 patients from six different institutions was selected providing comparable data sets with respect to one outcome measure, i.e. the Global Severity Index derived from the SCL 90-R (cf. Davies-Osterkamp, Strauss, & Schmitz, 1996).

To determine the relationship between the interpersonal problems at admission and the symptom related treatment outcome, criteria for clinically significant changes were determined (critical change value, $kb = x_{norm} + 2\_{norm}$; reliable change, $RC = (x_{post} - x_{pre})/s_{diff}$, i.e. standard error of the differences; cf. Jacobson & Truax, 1991).

Overall, the effect size for the SCL 90-R changes was .57, for changes in interpersonal problems (total IIP-score) it was .30. Using the criteria for clinically relevant improvement, four groups of patients could be distinguished:

- patients whose condition deteriorated (9.3%)
- patients with no perceptible changes (26.8%)
- patients whose symptoms improved (29.4%) and
- patients whose discharge values in the SCL-90-R fell into a normal range (20.2%)

An additional subgroup of 14.4% of the patients showed normal SCL-90-R scores at admission as well as at discharge („healthy“ subgroup).

In this study, the main goals were to discover whether patients who were more aware of their interpersonal problems before treatment began would have a better prognosis in inpatient group psychotherapy in terms of symptomatic change. It turned out that particularly those patients who were rated as "cured" or "improved" by the end of therapy on the basis of their SCL-90-R measures reported having the most interpersonal problems before their treatment began (cf. Fig. 2; $F[1,189]=6.7$; $p<.05$)

Insert figure 2 about here

Study 3

The overall objective of the multicentre study 3 was not only to replicate findings of the preceding studies on a larger scale, but also to further evaluate the differential and prognostic validity of different facets of interpersonal problems. The aim was to examine the prediction of outcome by interpersonal problems, thus addressing the "interpersonalness" of the treatment outcome. More specific research questions were to explore whether using ipsatised or unipsatised scores of the IIP would have any impact on this prediction (see below) and whether the interpersonal profile is related to changes in interpersonal problems and to personality related outcome.

Data for this study were collected within six different hospitals with different treatment settings (Kiel, Bad Honnef, Groenenbach, Goettingen, Geldern, Haldensleben). Again, the common element of the settings was the use of psychodynamic group work as a central part of each programme. Measures common to all settings were the IIP (German version, Horowitz, Strauss, & Kordy, 1994), the SCL-90-R (German version, Franke, 1992) and the Giessen Test (Beckmann, Braehler, & Richter, 1990), which is a psychoanalytically oriented personality inventory very commonly used in Germany. These measures were assessed at least twice, on admission and at discharge. In three of the six centres individual treatment goals were also formulated and recorded at the beginning of treatment. The extent (percentage) to which these goals were attained was assessed at the end of treatment by both the patient and the therapist. Finally, global ratings of somatic and psychological treatment outcome were given by the therapist at discharge.

For a better appreciation of the interpretation of the results, we should comment on some specific aspects of the instruments: The Giessen Personality Test has its roots in the German psychoanalytic and social psychology tradition.

However, it is less intended to capture individual traits, but rather conceptualises aspects of personality as they appear in dyadic relationships or in group processes. 40 items, using bipolar 7-point scales, include questions on emotional states, questions on ego qualities (such as introspection, imagination and permeability), questions on interpersonal states (such as self-disclosure) and questions on social reactions and/or responses by others. A factor analysis of the items yielded the following six scales:

The first scale is called “social response” and gauges the positive or negative feedback a person receives from his/her environment on the basis of his personality. The second scale is called “dominance” and assesses whether a person is more flexible, accommodating and obsequious or more domineering in social relationships. The scale “self-control” is less conceptualised from an interpersonal stance and addresses whether a person is more self-controlled, and even obsessive-compulsive, or more disorganised in his or her personal affairs. The fourth scale “underlying mood” represents a (hypo)manic-depressive dimension. The scale “permeability” assesses whether a person is more open, self-disclosing and permeable or more retentive and reserved. The sixth scale, “social potency” describes whether somebody is able to achieve his or her social goals in a maximally adaptive way.

The original factor analysis was based upon a sample of psychotherapy patients from the University Hospital Giessen. The 5-factor solution aacounted for 58% of the variance. Beckmann, Brähler & Richter (1990) report several replications of the factorial structure in different clinical and non-clinical samples.

The Giessen Test is not considered to be a genuine tool of psychotherapeutic outcome measurement. However, it has proved to be sensitive to change in specific contexts, i.e. in identifying psychotherapeutic treatment outcome (Stuhr, 1997), or changes in the self concepts of the German population during societal changes (Braehler & Richter, 2000). Because it was developed from a psychoanalytic and social psychology tradition, it has considerable face validity as a clinically relevant outcome measure for inpatients

With regard to the IIP, it is important to note that two different interpretations of the circumplex of interpersonal problems exist. These different lines of interpretation have emerged since interpersonal problems are considered to be the counterparts of interpersonal traits. In the tradition of circumplex models, these traits represent the vectors in a two-dimensional circular space formed by the

coordinates of affiliation (LOVE) and dominance (DOM). The eight scales (cf. Fig. 1) usually identify particular patterns of interpersonal tendencies. Wiggins, Phillips, & Trapnell (1989) as well as Gurtman (1991) have proposed to classify subjects according to the typological sector of the interpersonal circle in which they fall. This is defined by the average directional tendency of their interpersonal behaviour with reference to the coordinates of love and dominance. Ambiguous definitions of the interpersonal tendency have emerged when applying the interpersonal problems counterpart to this conceptual framework.

First, the typological sector of a person may be interpreted according to the content of the scale, i.e. these categories represent the interpersonal field that distresses the person to the highest extent. In this vein, one can imagine that a patient, who suffers from being too obedient and too submissive in interpersonal relationships, may also overtly complain about interpersonal disturbances related to assertiveness and dominance and not only to submissiveness.

The second interpretation is much more common and defines the category as an individual's interpersonal focus or tendency, i.e. an area that is the focus of one's inner attention and that leads to the most characteristic and precarious interpersonal conflicts. In line with this interpretation, Alden, Wiggins, and Pincus (1990) were able to show high convergent validity between interpersonal traits, as measured with the Revised Interpersonal Adjective Scales, and interpersonal problems. This finding indicates that it is rather the content of the scales or interpersonal areas which is evoked by the IIP items. A way to deal with these ambiguities, is to use unipsatised scores, when the literal meaning of the scales is more important and to use ipsatised scores, when the interpersonal tendency or focus has a higher validity. Ipsatizing is accomplished by subtracting an individual's mean from each IIP scale mean. Using the IIP in this large sample, we wished to elucidate these concurrent notions of the interpersonal tendency in a clinical population.

Following a replication of the findings from study 2, the analytic strategy behind this work was primarily to run multiple regression analyses to predict treatment outcome and secondly, to evaluate the impact of the interpersonal profile , on the eight octants respectively, on changes in interpersonal and personality related measures. These research questions are part of a larger project which also addresses the relationship between interpersonal problems and conventions of clinical and statistical significance on the one hand and interpersonal problems and diagnostic categories, especially personality disorders on the other. Thus, we

present selected results related to the clinical validity of the Inventory of Interpersonal Problems.

Sample characteristics

Complete data for the IIP and SCL-90-R were obtained from about 740 patients at both points of measurement. Approximately 600 patients completed the Giessen Test two times. However, the goal attainment and global outcome ratings were only assessed in three hospitals. 350 global outcome ratings were collected in these centres; 260 patients and 120 therapists assessed the extent to which the initial treatment goals had been reached.

67 % of the subsample of patients who completed the questionnaires at admission and at discharge were female and 33 % were male. The mean age of these patients was 38 (+/- 9.5; range: 17-63). The global diagnoses represented in this sample were personality disorders (22%), eating disorders (21%), anxiety, dissociative and somatoform disorders (22%), affective disorders (19%), substance abuse (5%) and psychoses (0,5%).

Replications of findings on the relationship between clinically significant change and interpersonal problems

Applying the conventions of clinical and statistical significance, 25 % of the sample fell into the “healthy” or “normal range” category, 28% improved in clinically and statistically significant terms (“cured”) and 15 % remained unchanged. 14% of all patients were assigned to the category “deteriorated” (statistically and/or clinically significant). These proportions are very similar to those identified in the second study.

Again, patients who deteriorated, showed the lowest level of interpersonal problems and the least differentiated interpersonal profile. It should be noted, however, that some contrasting findings related to this group were obtained from the global outcome ratings and goal attainment scaling of the therapists. Contrary to expectation, the outcome ratings of the therapists were highest in patients who deteriorated in the SCL90R! The patients who deteriorated according to the SCL-90 change scores were also the only group that had higher scores on all interpersonal problem scales at discharge no matter what area they were related to, as shown in figure 3.

Figure 3: Interpersonal problems on admission and at discharge in patients who deteriorated statistically and clinically

The interpersonal profile of the IIP pre- and post-treatment scores shown in Fig. 3 sharply contrasts with the profile of patients who improved (Fig. 4). The patients who improved according to statistical and clinical criteria complained less about interpersonal problems, especially about those located in the lower octants (too submissive), while problems with being overly autocratic were reported more often.

The fact that patients who deteriorated in clinically and statistically significant terms reported only few interpersonal problems on admission, yet more problems at discharge could have led their therapists to globally rate a positive treatment outcome. A more detailed discussion of these findings will be provided by Strauss, Schmidt and the working group on inpatient group therapy (in preparation).

Insert Figure 4 about here

Figure 4: Interpersonal problems on admission and at discharge in patients who improved statistically and clinically

Prediction of treatment outcome on the basis of the quality of interpersonal problems

To test if the quality of interpersonal problems (i. e. the pattern of scores in the eight subscales) can predict outcome, multiple regression analyses were performed in two steps: We first included the unipsatised scores, then in a second step the ipsatised scores as predictors into the equation. In both equations the mean score of the IIP was included as well as a predictor variable in order to explore whether the total amount of interpersonal problems or specific aspects have a higher prognostic impact. Criterion variables were a) the difference score on the Global Severity Index (GSI), b) the difference scores on the scales of the Giessen Test and c) the global outcome ratings as well as the goal attainment scales. Global Severity Index

Because of the nature of the circumplex, a high degree of multicollinearity has to be taken into account when interpreting findings. The correlations between the IIP total score and the single scales was highest in HI (too unassertive), FG (too socially avoidant), and JK (too exploitable). Thus, the predictive power of one of these scales might easily be suppressed by the others.

Predictions of goal attainment and global outcome ratings

In predicting the goal attainment ratings and the global outcome of both patient and therapist, interpersonal problems had a significant effect only on the therapists` global ratings of the psychological outcome (R^2 =.7, p <.001). In this regression, nonassertiveness at admission had a significant impact on a better global psychological treatment outcome as rated by the therapist. This is consistent with the findings of Horowitz, Rosenberg, and Bartholomew (1993). They concluded that problems with friendly submissiveness seem to be more easily treated in brief dynamic psychotherapy than problems with hostile dominance.

Prediction of improvement in the Global Severity Index (GSI)

In predicting the change in the GSI, both the regression equation including ipsatised scores as well as the equation including unipsatised scores were found to be significant. Interpersonal problems explained approximately 12 % of the variance of change on the GSI. In both equations, only the IIP total score was found to have a significant effect on change in the GSI scores (F = 19.60; p <.001), but none of the single scales.

Prediction of change on the scales of the personality inventory (Giessen Test)

In predicting change on the scales of the Giessen Test, the single scales had a greater predictive potential than the total IIP scores (Tables 1-5).

Insert tables 1-5 about here

Table 1: Prediction of „social response“ by IIP scales[1]

	β of unipsatized scale scores	β of ipsatized scale scores
PA	-.06	.02
BC	-.10*	-.05

DE	-.09	-.03
FG	-.04	.02
HI	-.12**	-.10*
JK	-.15*	-.06
LM	-.05	.01
NO	.01	.06
Mean	-.07	-.24**
R	.33	.27
R^2	.10	.08
F	17.30	22.92
P<	.001	.001

[1]Note that all regression equations employed the stepwise procedure; the solution presented is based on the final model; all regression coefficients are standardised

Table 2: Prediction of „dominance“ by IIP scales

	β of unipsatized scale scores	β of ipsatized scale scores
PA	-.32**	-.31**
BC	-.02	-.01
DE	-.07	-.15**
FG	.05	.06
HI	.13*	.00
JK	.03	-.04
LM	.03	-.02
NO	-.01	-.10*
Mean	-.01	-.16**
R	.33	.37
R^2	.10	.12
F	19.66	17.68
P<	.001	.001

Table 3: Prediction of „underlying mood (depression)“ by IIP scales

	β of unipsatized scale scores	β of ipsatized scale scores
PA	-.06	-.06
BC	.03	.02

DE	-.12	-.08
FG	-.08	-.06
HI	.04	.03
JK	.09	.06
LM	.12	.08
NO	.01	.01
Mean	.22**	.23**
R	.17	.27
R^2	.05	.08
F	27.25	27.25
P<	.001	.001

Table 4: Prediction of „permeability“ by IIP scales

	β of unipsatized scale scores	β of ipsatized scale scores
PA	-.01	-.06
BC	.08	.12*
DE	.05	.03
FG	.01	.01
HI	-.01	.02
JK	.02	.08
LM	.17**	.16**
NO	.04	-.04
Mean	.03	.16**
R	.16	.23
R^2	.04	.06
F	19.66	6.84
P<	.001	.001

Table 5: Prediction of „social impotency“ by IIP scales

	β of unipsatized scale scores	β of ipsatized scale scores
PA	-.04	-.12*
BC	.04	.01
DE	.10	.02
FG	.19**	.02
HI	.09	.02
JK	.06	.01
LM	.06	.02
NO	-.10	-.13*
Mean	.07	.15*
R	.19	.21
R^2	.04	.05
F	18.71	12.00
P<	.001	.001

In general, all equations predicting change in the scales of the Giessen Test, based on interpersonal problems, were significant except for change in the self-control scale. Another general finding was that a decrease in the depression scale was predicted by the IIP total score, but neither of the single subscales proved to be significant. However, the single IIP scales had a significant impact on changes in the four Giessen Test subscales, which more strongly reflect relational dimensions.

The prognostic effects of the IIP scales were overall highly significant, irrespective of the fact that ipsatised or unipsatised scores were used as predictors. Because of the high multicollinearity of the IIP scales, it is difficult to compare the effect of using ipsatised and using unipsatised scores in the regression equations.

The impact of the interpersonal focus on changes in the Giessen Test

In addition to the regression analysis, a finer analysis of the validity of the interpersonal focus was performed by analysing the effect of the average interpersonal tendency as suggested by Wiggins, Phillips and Trapnell (1989) and Gurtman (1991). According to this suggestion, the sample is divided into subgroups of patients showing their predominant interpersonal distress within the different octants of the circumplex model. The categories to which the patients were assigned were roughly equally distributed (n=93 – 110). In table 6, the

mean changes related to aspects of personality, as measured by the Giessen Test, are shown across these interpersonal categories.

Table 6: Interpersonal focus and changes in the Giessen test

	Interpersonal focus in the twodimensional space										
Changes in scales of the Giessen test	**PA**	**BC**	**DE**	**FG**	**HI**	**JK**	**LM**	**NO**	**mean**	**F**	**p**
Social Response[1]	-.49	-1.33	-1.71	-2.71	-2.51	-2.40	-1.31	1.79	-1.42	2.38	.021
Dominance[2]	-1.10	-2.45	.67	2.11	2.72	2.84	.65	-.63	.67	6.12	.000
Self-control[3]	-.98	-.07	.75	-1.01	.40	.45	.18	-.41	-.12	.85	.547
Mood (Depression)[4]	3.25	3.48	3.14	3.38	5.24	4.42	2.78	2.22	3.50	1.23	.284
Permeability[5]	1.09	1.03	.11	.95	2.04	2.62	1.38	-1.49	1.07	.89	.533
Social Potency[6]	2.18	1.22	3.49	5.98	4.80	2.76	1.96	1.71	3.03	2.85	.006

[1] Negative signs indicate that patients received more positive feedback from others at discharge.

[2] Negativ signs indicate that patients were more flexible at discharge.

[3] Negative signs indicate that patients had a higher self-control at discharge.

[4] Positive signs indicate a decrease of the depressive mood.

[5] Positive signs indicate that patients were less retentive and more self-disclosing or permeable at discharge.

[6] *Positive signs indicate that patients were more socially potent at discharge.*

A descriptive analysis of the relationship between these categories and differences in the Giessen Test revealed that all groups improved on the depression and social potency scales. To test differences between the categories, univariate analyses of variance were performed (table 6). On the social potency scale, the mean differences were found to be significant within the categories BC (vindictive) and NO (intrusive) showing lower scores than FG (socially avoidant) and HI (unassertive).

All other scales showed differential effects in the eight categories. Partially, these differences can be attributed to the fact that patients with an interpersonal focus on FG, HI and JK (socially avoidant, unassertive and exploitable) obtained the highest scores on admission. However, some peculiar findings occurred. The mean differences were significant on the social response scale revealing a higher social response in patients from all categories, in particular FG, HI, JK, but a lower social response in the NO (intrusive) category. This was the only category with lower scores on permeability indicating that patients assigned to this group were more reserved after treatment. The differences in this scale were, however, not statistically significant. Another finding, which was theoretically consistent, was that in patients from all categories of the upper octants, dominance as a personality trait decreased while it increased in patients located in two of the lower octants. These findings underline those of the regression analyses indicating that the interpersonal focus clarifies the spectrum of the deeper dimension of personality related changes.

The impact of the interpersonal focus on changes in interpersonal problems

The interpersonal focus should not only be important in predicting various types of outcome , but also in predicting general change in interpersonal problems during treatment. Table 7 provides the change scores of interpersonal problems across the eight interpersonal categories.

Table7: Interpersonal focus and changes in interpersonal problems

	Interpersonal focus in the twodimensional space									
Changes of IIP scales*	PA	BC	DE	FG	HI	**JK**	LM	NO	F	p
Change in PA	***.24***	.24	.12	.05	-.17	-.14	.12	.22	8.76	.001
Change in BC	.07	***.12***	.13	.13	.00	.00	.02	.05	2.60	.01
Change in DE	.00	-08.	***.27***	.23	.12	.03	.02	.05	10.18	.001
Change in FG	.10	.26	*.51*	***.73***	***.63***	.53	.12	.18	12.23	.001
Change in HI	.00	.00	.34	.55	***.61***	*.72*	.38	.11	10.88	.001
Change in JK	.11	.00	.00	.27	.51	***.64***	.42	.11	5.40	.001

Change in LM	.21	.00	.12	.23	.38	.54	***.48***	.23	7.15	.001
Change in NO	.25	.16	-.05	.03	.02	.25	.46	***.36***	4.90	.001
Total change	.15	.15	.23	.29	.29	.31	.22	.15	4.75	.001

* The scores are based on unipsatized scores.

Patients with an interpersonal focus in the lower octants showed the greatest decreases of interpersonal problems during treatment pertaining to the fact that their baseline scores were especially high (regression to the mean). The HI and JK categories displayed even more interpersonal disturbance related to being too autocratic at discharge. One would assume that changes in interpersonal problems should be highest in the interpersonal area that a patient most sensitively perceives. It was possible to confirm this assumption in most categories. There were only two exceptions: patients from the categories BC and DE did not show any obvious alterations in their most conflictual interpersonal problem areas. Patients with a focus on BC showed no changes in a couple of interpersonal dimensions (for instance DE, HI, JK, LM). One might conclude that patients who have their interpersonal focus in the cold and vindictive area improve in terms of intrapsychic aspects as measured in the Giessen Test, however they do not change their dismissing interpersonal stance.

Discussion

The purpose of this article was to describe research activities in a specific field of psychodynamic psychotherapy, i.e. inpatient group work. Psychodynamic group treatment is the core element of many treatment programmes within psychotherapeutic hospitals at least in Germany. Nevertheless, it has been somewhat neglected in research. This was the reason to found a working group on inpatient group psychotherapy aiming to promote research within this field in several naturalistic studies. The studies summarised in this article mainly focussed on the question as to whether interpersonal characteristics of patients treated in inpatient groups might be used to differentially predict treatment outcome on different levels.

One general result of the three studies performed by the working group is that patients who are more susceptible to interpersonal problems at the beginning of their treatment (i.e. patients with higher total scores in the IIP) seem to have more benefit from the psychodynamic treatment than patients suffering less from interpersonal distress. This benefit is related to symptomatic changes as well as personality-related outcome measures (Giessen Test). In some ways, this result could have been expected since psychodynamic group treatment should specifically focus on interpersonal issues. Accordingly, patients who are already familiar with these problems (perhaps related to “psychological mindedness”) should have a better outcome. Nevertheless, even expected results should be considered in planning treatment programmes. One potential conclusion that could be drawn from this result might be to assign patients with lower scores in the IIP to other treatments or - at least – to provide additional treatments focussing on increasing interpersonal sensitivity.

Apart from this global result, differential analyses of the data from the recent study give some additional information regarding the patients benefit from psychodynamic group work in terms of personality related criteria: Being too nonassertive, too socially avoidant or too vindictive at admission increases the likelihood of receiving more positive feedback from the environment at the end of the treatment (as measured by the Giessen Test scale “social response”). Being too domineering at admission had a significant impact on being more flexible and accommodating at the end of treatment, whilst being too nonassertive on admission had a significant effect on being more dominant (Giessen Test scale “dominance”). Using the ipsatised scores in this regression, being too domineering, too intrusive and too vindictive was highly predictive of being more flexible and less dominant at the end of treatment. Interpersonal

problems related to being too nurturant were predictive of higher permeability or self-disclosure at the end of treatment. This finding is not consistent with expectations. In line with the assumptions based upon interpersonal theory we would have expected patients whose interpersonal focus is related to nurturance, to benefit from therapy by being more reserved at discharge. The finding that patients who complained about being too vindictive on admission were at the time of discharge more permeable or self-disclosing was, however, in line with our assumptions. Being too socially avoidant predicted higher social potency or an increased capacity to carry out one's own desires at the end of treatment. It is important to note that, using the ipsatised scores, interpersonal problems of being too domineering and too intrusive predicted higher *social impotence*.

Summarising these findings, the significant contributions of the IIP scales in predicting change in the personality test underlines not only the construct validity of interpersonal problems (cf. Davies-Osterkamp & Kriebel, 1993), but also in particular its predictive validity. Both different notions of the interpersonal problems scales seem to have their own predictive potential. In general, the use of ipsatised scores more strongly emphasises the upper octants while the use of unipsatised scores increases the impact of the lower octants, probably because the scores on these scales showed the highest correlations with the IIP total score. One might conclude that using the unipsatised scores more strongly emphasises the effects on interpersonal problems in their literal meaning, whilst the use of ipsatised scores stresses more the interpersonal focus or tendency.

On the whole, the finding of Horowitz et al. (1993) that problems with friendly submissiveness seem to be more easily treated in dynamic psychotherapy than problems with hostile dominance, was not obtained in the setting of inpatient psychodynamic group therapy. Even though one might reach the conclusion that patients with a focus on friendly submissiveness might benefit more from group therapy, when simply looking at the unipsatised effects of the IIP scales, this assumption cannot be confirmed on the basis of the ipsatised scores. There was no kind of interpersonal problem dimension that did not have any positive effect on treatment outcome. We interpret these results to conclude, that in group therapy it might be possible to take specifically targeted action regardless of what kind of interpersonal problems patients experienced predominantly at the beginning of therapy. These actions were aimed at achieving a new perception or a more positive stance related to specific interpersonal

problems in each individual of the group. The mutual interactions in group therapy may specifically offer a better reflection of interpersonal aspects of the self and others.

References

Alden, L.E., Wiggins, J.S, & Pincus, A.L. (1990). Construction of circumplex scales for the Inventory of Interpersonal Problems. *Journal of Personality Assessment, 55:* 521-436.

Beckmann, D., Braehler, E., & Richter, H.E. (Eds.) (1990). *Der Giessen-Test (GT): Handbuch* (The Giessen Test: manual). Bern: Hans Huber.

Braehler, E., & Richter, H.E. (2000). Das psychologische Selbstbild der Deutschen zur Jahrtausendwende (The psychological self image of Germans in the new millenium). In O. Decker & E. Braehler (Eds.) *Deutsche zehn Jahre nach der Wende* (pp.47-51). Giessen: Psychosozial Verlag.

Davies-Osterkamp, S., & Kriebel, H. (1993). Konstruktvalidierung von Symptomskalen und Persönlichkeitsstests durch das Inventar zur Erfassung Interpersonal Probleme (Validating constructs of symptom scales and personality tests with the Inventory of Intpersonal Problems). *Gruppenpsychotherapie und Gruppendynamik, 29*: 295-307.

Davies-Osterkamp, S., Strauss, B., & Schmitz, N. (1996). Interpersonal problems as predictors of symptom-related treatment outcome in long term psychotheapy. *Psychotherapy Research, 6*: 164-176.

Eckert, J., Biermann-Ratjen, E., Brodbeck, D., Burgmeier-Lohse, M., Keller, W., Schulz, E., Schuricht, C., & Strauss, B. (1997). Indikation für Psychotherapie: Welchen Einfluss nehmen interpersonale Probleme des Patienten auf die Indikationsstellung und die Wahl des Settings? (Treatment selection: Which influences do interpersonal problems of patients have on the assignment to psychotherapy and the choice of the treatment setting?). *Gruppenpsychotherapie und Gruppendynamik, 33*: 1-17.

Fonagy, P. (1999) (Ed.). *An open door review of outcome studies in psychoanalysis*. London: University College.

Gurtman, M.B. (1991). Trust, distrust and interpersonal problems: A circumplex analysis. *Journal of Personality and Social Psychology, 62*: 989-1002.

Franke, G.H. (1992). Eine weitere Überprüfung der Symptom-Check-Liste (SCL-90-R) als Forschungsinstrument (A further evaluation of the symptom check list (SCL-90-R) as a research tool). *Diagnostica, 38*: 160-167.

Horney, K. (1945). *Our inner conflicts*. New York. Norton.

Horowitz, L.M., Rosenberg, S.E., Baer, B. A., Ureno, G., & Villasenor, V.S. (1988). Inventory of interpersonal problems. *Journal of Consulting and Clinical psychology, 56*: 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, 69:* 549-590.

Horowitz, L.M., Strauss, B., & Kordy, H. (Eds.) (1994) *Manual zum Inventar zur Erfassung für interpersonale Probleme* (IIP-D) (Manual of the Inventory of Interpersonal Problems, German version). Weinheim: Beltz.

Jacobson, N.S. & Truax, P. 81991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical psychology, 59:* 12-19.

Mohr, D.C:, Beutler, L. E., Engle, D., Shoham-Salomon, V., Bergan, J., Kaszniak, A.W., & Yost, E.B: (1990). Identification of patients at risk for nonresponse and negative outcome in psychotherapy. *Journal Consulting and Clinical psychology, 58*: 622-628.

Muhs, A. (1993). Results from the psychosomatic hospital Mannheim. *Gruppenpsychotherapie und Gruppendynamik, 29*: 259-270.

Rad, M. v., Senf, W., & Bräutigam, W. (1998). Psychotherapie und Psychoanalyse in der Krankenversorgung: Ergebnisse des Heidelberger Katamneseprojekts (Psychotherapie and psychoanalysis in the health system – Results from the Heidelberg follow-up study). *PPmP, Psychotherapie, Psychosomatik, Medizinische Psychologie, 48*: 88-100.

Strauss, B. (1992). Empirische Untersuchungen zur stationären Gruppenpsychotherapie (Empirical studies on inpatient group psychotherapy). *Gruppenpsychotherapie und Gruppendynamik, 28:* 125-149.

Strauss, B. & Burgmeier-Lohse, M. (1993). Ergebnisse aus der Klinik für Psychotherapie Kiel (Results from the University Hospital Kiel). *Gruppenpsychotherapie und Gruppendynamik, 29*: 237-244.

Strauss, B. & Burgmeier-Lohse, M. (1994). *Stationäre Langzeitgruppen-psychotherapie* (Inpatient longterm group psychotherapy). Heidelberg: Asanger.

Strauss, B. & Burgmeier-Lohse, M. (1995). Merkmale der Passung zwischen Therapeut und Patient als Determinante des Behandlungsergebnisses in der stationären Gruppenpsychotherapie (Characteristics of the patient-treatment-fit as

a determinant of treatment oucome in inpatient group psychotherapy). *Zeitschrift für Psychosomatische Medizin und Psychoanalyse, 41:* 127-140.

Strauss, B., & Kaechele, H. (1998). The writing on the wall: Comments on the current discussion about empirically validated treatments in Germany. *Psychotherapy Research, 8:* 158-170.

Strauss, B., Eckert, J., & Ott, J. (1993) (Eds.), Interpersonale Probleme in der stationären Gruppenpsychotherapie – Themenheft (Interpersonal problems in inpatient psychotherapy – Special issue). *Gruppenpsychotherapie und Gruppendynamik, 29(3):* 223-323.

Strauss, B., Eckert, J., & Hess, H. (1993). Integration und Diskussion der Ergebnisse (Integration and discussion of the results). *Gruppenpsychotherapie und Gruppendynamik, 29*: 286-294.
Strauss, B., Kriebel, R., & Mattke, D. (1998). Probleme der Qualitätssicherung in der stationären Gruppenpsychotherapie (Problems of quality assurance in inpatient group psychotherapy). *Psychotherapeut, 43*: 18-25.

Strauss, B., Lobo-Drost, A., & Pilkonis, P.A. (1999). Einschätzung von Bindungsstilen bei Erwachsenen – erste Erfahrungen mit der deutschen Version einer Prototypenbeurteilung (Assessment of attachment styles in adults – first experiences with the German version of a prototype rating). *Zeitschrift für Klinische Psychologie, Psychiatrie und Psychotherapie, 47*: 347-364.

Strupp, H.H., Horowitz, L.M, & Lambert, M. J. (1997) (Eds.). *Measuring patient changes*. Washington: American Psychological Association.

Stuhr, E. (Ed.) (1997) *Therapieerfolg als Prozess- Leitlinien fuer eine kuenftige Psychotherapieforschung*. (Therapeutic success as process. Guidelines for future psychotherapy research. Asanger: Heidelberg.

Tschuschke, V. (1993). *Wirkfaktoren stationärer Gruppenpsychotherapie* (Therapeutic factors of inpatient group psychotherapy). Göttingen: Vandenhoek & Ruprecht.

Wiggins, J.S., Phillips, N., & Trapnell, P. (1989). Circular reasoning about interpersonal behaviour: Evidence concerning some untested assumptions underlying diagnostic classification. *Journal of Personality and Social Psychology, 56*: 296-30

Investigating structural change in the process and outcome of psychoanalytic treatment - The Heidelberg-Berlin Study

Tilman Grande, Gerd Rudolf, Claudia Oberbracht,
Thorsten Jakobsen, Wolfram Keller

Psychosomatische Klinik
Thibautstraße 2
69115 Heidelberg

European Studies in Psychoanalytic Psychotherapy

Germany

1. History, Focus and Present State of the Research Project

In 1993 the Confederation of German Psychoanalysts (Deutsche Gesellschaft für Psychoanalyse, Psychotherapie, Psychosomatik und Tiefenpsychologie - DGPT) decided to support research on psychoanalytic long-term therapy. The DGPT's plan was to study high intensity long-term psychoanalyses in comparison with one hour per week psychodynamic psychotherapies. In calling for research proposals the DGPT's essential objective was to subject the effectiveness and efficiency of long-term therapies to empirical investigation and thus marshal arguments that could be deployed in the ongoing health-policy debate in Germany. However, its initiative also coincided with an existing interest on the part of psychoanalytic organizations to evaluate their own work and gain a deeper understanding of it with the aid of systematic research. Aside from the question of effectiveness, the call for projects was thus also guided by an internal interest in the investigation of the processes and specific change mechanisms activated by psychoanalytic treatments.

In taking up the DGPT's invitation, our considerations soon centred on the fundamental question of whether, in view of the present state of research and methodology, their demand for a research project on the psychoanalyses could in fact be satisfied. The most important problem related to the fact that, the research instruments used so far, against the background of the so-called dose-effect models (Howard, Kopta, Krause & Orlinski 1986), had identified significant changes in the early phases of psychotherapeutic treatment but only minor effects in the further course of treatment. On the basis of studies indebted to this model, Grawe, Donati und Bernauer (1994) queried the usefulness of long-term therapies and concluded that positive effects can only be expected from treatments of up to about 50 hours.

Our position on this is a very different one. We are convinced that the specific effects of psychoanalyses, which, as clinical experience shows, only materialize after long and intensive treatment, cannot be detected with the aid of conventional research instruments. These measuring instruments operate close to the surface, capturing above all symptomatic or behavioural characteristics (Grande & Jakobsen 1998) whereas from a psychoanalytic point of view, the essential changes take place at the level of personality structure, i.e. in the course of the breakdown of pathological structures that have taken shape in the course of an individual’s development and the reorganization and/or reintegration of the pathogenic intrapsychic conflicts and vulnerabilities embodied in those structures. Such processes of restructuring, which in all probability can only be achieved by means of long-term analytical processes, can obviously not be registered by conventional change-assessment techniques (cf. Strupp, Schacht & Henry 1988). These considerations led us to conclude that a project dealing with the effectiveness of long-term psychoanalytic therapies could only be successful if a method were available to assess central personality structure from a psychoanalytic point of view.

To this end a research group was constituted in 1992 involving 40 scientists and clinicians with a psychoanalytic background and from 12 different universities; within the framework of the project on an Operationalized Psychodynamic Diagnosis system (OPD), the group developed instruments designed to remedy this situation (Arbeitsgruppe OPD 1998). Four years work led to a classification system for research, teaching and practice based on psychoanalytic constructs and thus transcending the purely descriptive approach underlying the existing systems (ICD, DSM). The OPD instruments thus assume a central position in our research approach, which, towards the end of 1993, was presented to the DGPT in the form of a project proposal (Rudolf & Grande 1997; Grande, Rudolf & Oberbracht 1997). The plan was evaluated by several independent experts and classified as suitable for responding to the issues posed by the DGPT's call for projects.

In 1997, after a preliminary study, we were able to begin work on the project and start collecting data in Heidelberg (research group: Rudolf, Grande, Oberbracht, Jakobsen) and Berlin (Keller, Dilg, Stehle). By mid-2000 a total of 61 cases from Heidelberg and Berlin had been incorporated into the study; the average length of therapy was approx. 2 years. As recruitment is still in progress and the therapies are long-term in design, the study itself is not yet complete. Accordingly, the present article presents a discussion of the study design and the methods we have developed to measure structural changes. We will exemplify this procedure with reference to observations made in the actual course of the psychoanalytic treatment of one particular patient. Further, we report on the outcome of preliminary investigations (now complete) on the reliability and validity of this method.

Central to our investigation is the question of the specific quality of therapeutic changes that take place in intensive psychoanalytic treatment on the one hand, and in low-frequency psychodynamic therapies on the other. Our hypothesis is that there are qualitatively distinct forms of change that can be designated as "coping" in the one instance and "structural change" in the other. We assume that "structural changes" are more likely to occur in psychoanalyses, whereas in low-frequency therapies the probability is higher of encountering changes of a "coping" kind. We also hope toshow how these two types of change have been conceptualized and operationalized in the study. In the final section of the article we return to this distinction and refine our hypothesis with reference to the investigatory methods we have employed.

2. Study design, sample and research instruments

Figure. 1 shows the study design. It involves 30 patients receiving psychoanalytic psychotherapy of at least 3 hours a week and 30 psychotherapy patients in a one-hour sedentary setting. To protect this number of cases against attrition, 36 patients were admitted to each group. At the outset of therapy the patients are subjected to detailed diagnostic examination. In the further course of the study assessments are made at 3- or 6-monthly intervals through to treatment termination. Recent research by Sandell, Blomberg & Lazar (1999) strongly suggests that frequency and setting differences have their clearest effects in the years following the termination of therapy. Accordingly a further central line of inquiry examines the changes taking place in that period; hence the research design also extends to follow-up studies. Follow-up studies are to take place one and three years after the termination of therapy.

Fig. 1: Study Design

The patients admitted to the study display severe neurotic, psychosomatic and personality disorders. In our study patients are classified as severely disordered if they present a moderate or low integration level on the “Structure” axis of the OPD (see Section 3.1 below) and in addition display clear-cut symptoms measured by an Impairment Severity Score (Schepank 1995). The patients are assessed on the basis of these criteria before admission. With an eye to ensuring continuous comparability of the differential effects in the two groups, each of the

participating analysts brings to the project one psychoanalytic case and one psychotherapy case. In this way possible confounding factors associated with the person of the analyst are compensated for in both groups.

Both treatment groups are matched for sex, age, educational level and disorder severity (usingthe above-mentioned inclusion criteria) so that some level of patient comparability is also assured. The alternative – of employingrandomized group allocation - is hardly feasible under normal outpatient conditions and carries its own substantial scientific risks regarding the validity of the investigation (by distorting the object of study). Hence a decision was made in favour of a naturalistic design plus subsequent group matching. In many other details of the study care is also taken to ensure protection of the therapeutic situation and as little interference from our research as possible. For example, the data collection procedure is organized in such a way that analyst and patient negotiate involvement in the study before the onset of therapy but not in the course of it. Where a degree of research impact on therapeutic work is unavoidable this is documented both by the external raters and the analysts and incorporated in the evaluation.

The data collected in the course and at the end of therapy stem from four different observational vantage points: patient self-assessment, analyst assessment, health insurance data, and assessments by external raters (cf. Grande, Rudolf & Oberbracht 1997). The following is an (incomplete) overview of the instruments used and the domains captured by these assessments.

1. ***Patient self-assessment*** at the beginning and in the course of therapy; the following data are collected: social status / socio-demographic data / health and health behavior (items from the Berlin Jung study and the Berlin psychotherapy study, cf. Keller, Dilg, Westhoff et al. 1997, Rudolf 1991a) / SCL90-R (Derogatis, Lipman & Covi 1975) / Psychic and Social-Communicative Questionnaire PSKB-Se (Rudolf 1991b) / Inventory of Interpersonal Problems IIP (Horowitz, Strauß & Kordy 1993) / Health Scale (SG-Scale from the TPF; Becker 1989) / INTREX-Introject questionnaire (Benjamin 1974; Tress 1993).
2. ***Analyst assessment at the outset and termination:*** ICD-10 diagnoses (Dilling, Mombour & Schmidt 1991) / conflicts, structural level, severity of impairment (short version of OPD) / Initial Working Alliance iTAB (Grande, Porsch & Rudolf 1988; Rudolf 1991a) / physical symptoms and psychic symptoms (Heidelberg Documentation System, Rudolf, Laszig & Henningsen 1997).

   Analyst assessment in the course of therapy: information on setting and interruptions / significance of and changes in symptomatology over the preceding 3 months (free description) / analytic process and cooperation, contents and themes in the therapeutic work (free description) / session

protocols / Working Alliance TAB (Grande, Porsch & Rudolf 1988; Rudolf 1991a).

3. ***Health insurance data:*** number of days in hospital, sick leave, use of medical services 3 years prior to therapy and 3 years after termination of therapy; data provided by the health insurance institutions.
4. ***Assessment by external raters:*** OPD rating, focus selection and HSCS rating in the course of treatment based on videotaped interviews (explained in detail below).

The last of theselines of inquiry is a special feature of our study, given that the use of external observers is unusual in connection with psychoanalytic therapies. Practically all other existing studies limit themselves to assessments by patients and/or analysts, thus foregoing the independent and hence more objective judgment of external observers. The rationale for this is based on apprehension that the psychoanalytic process might be impaired by third-person involvement. Some studies draw on tape recordings of sessions but these cannot provide information matching the quality of clinical interviews where the degree of change achieved is subjected to systematic verification. In our study, clinical interviews carried out at regular intervals in the course of treatment by external assessors supply the material for the subsequent assessment of process and outcome by independent raters. This feature sets the study apart from existing empirical psychoanalytis research and represents one of its main methodological specialties.

The number of patients (N=30 for each group) may appear small in terms of statistical power compared with the methods usually used to measure therapeutic outcome and differential treatment effects. But the drawbacks of a small number of cases can be compensated for by specific hypotheses concerning the differences between the two forms of therapy and the use of instruments sensitive enough to capture those differences. In the study this is achieved by the use of instruments specifically designed for the psychoanalytic approach, notably the Structural Change Scale (see below).

In addition, restricting the study to a relatively small number of cases means that the use of time-consuming study procedures becomes a viable proposition. For example, the rating of one single videotaped interview takes 16 hours on average; the number of interviews per analysis can be anything from 6 to 12 or more, depending on the length of the treatment. No less time-consuming are the qualitative text evaluation procedures used to study the written reports of the analysts on the course of therapy. The material thus gathered is so extensive and detailed as to make individual case studies feasible over and above the findings on the groups.

3. An OPD-based procedure for measuring change

3.1. Operationalized Psychodynamic Diagnosis system (OPD)

We have already indicated the major significance of the methods developed by the *Operationalized Psychodynamic Diagnosis* system (OPD) Study Group for the investigation of the process and the effects of psychoanalytic treatment. The OPD (Arbeitsgruppe OPD 1998) encompasses *five axes*, of which three are *psychodynamic* in the stricter sense of the term: those pertaining to "Relationship" (axis II), "Conflict" (III), and "Structure" (IV). The other two axes assess illness related behavior with respect to the preconditions of treatment (motivation, resources, etc.) and psychiatric disorders in accordance with the *International Classification of Diseases* (ICD-10; Dilling, Mombour & Schmidt 1991). For the study design envisaged here, only the three psychodynamic axes are relevant.

- *The OPD Relationship Axis*: The *Relationship Diagnosis* identifies the core dysfunctional relationship pattern displayed by a patient. Integral to this pattern are interpersonal behaviours and positions taken up by the *patient* and his/her *objects* in the core problematic relational constellation repeatedly established by the patient. The specific quality of these positions and the relational behaviour associated with them is described for each patient individually with reference to a given list of 30 items.
  The first step is to map the pattern from the *subjective experiential perspective of the patient*. Here we draw on relationship episodes reported by the patient in the course of the interview - these allow conclusions about the internal ideas and perceptions entertained by the patient in connection with the problems he/she has in handling interpersonal relationships. – The second stage is to describe the pattern from the *subjective experiential perspective of others* (including the interviewer). Here we draw additionally on the relational behaviour of the patient in the interview itself, with the countertransference of the interviewer figuring as a source of information. The third stage integrates the two experiential viewpoints into a single *relation-dynamic formulation* combining the results of the first two stages to form a cogent relational gestalt (see Grande, Burgmeier-Lohse, Cierpka et al. 1997 for more details).
- *The OPD Conflict Axis*: Within the framework of *Conflict Diagnosis*, 8 conflict types are defined as having a potentially crucial effect on the lives of the patients. The scaling is used to assess how significant each of these conflicts is for the individual patient. The following types of conflict are encompassed: *dependency versus autonomy*, *submission versus control*, *need for care versus self-sufficiency*, *self-esteem conflicts*, *superego and guilt conflicts*, *oedipal-sexual conflicts*, *identity conflicts*; also included is the clinical syndrome described as *deficient awareness of feelings and conflicts*. The manual describes criteria for the elaboration of these conflicts in the

following areas: partner selection, attachment behaviour/family life, family of origin, behaviour in the vocational/professional sphere, behaviour in the socio-cultural environment, and illness behaviour. A four-tier scale is used to assess whether and with what degree of intensity a conflict is present. In addition, raters are instructed to indicate which two of these conflicts are most important for the patients. A concluding assessment records whether the patient's handling of the conflicts corresponds to a more active or passive mode.

Table 1: OPD axes and focus list

Relationship Individualized formulation of a core dysfunctional relationship pattern	
Life-determining conflicts 1. dependence/autonomy conflict 2. submission/control conflict 3. care/self-sufficiency conflict 4. self-value conflicts	 5. guilt conflicts 6. oedipal-sexual conflicts 7. identity conflicts 8. deficient awareness of feelings and conflicts
Structural capacities/vulnerabilities 1. capacity for experience of self *self-reflection* *image of self* *identity* *differentiation of affects* 2. capacity for self-regulation *affect-tolerance* *regulation of self-esteem* *regulation of impulses* *anticipation* 3. capacity for defence *intrapsychic defenses* *flexibility*	 4. capacity for object-experience *self-object differentiation* *empathy* *awareness of total objects* *object-related affects* 5. capacity for communication *contact* *decoding other's affects* *encoding own affects* *reciprocity* 6. capacity for attachment *internalizations* *detaching* *variability of relationships*

- *The OPD Structure Axis*: In *Structure Diagnosis*, the patient's level of functioning and integration is assessed on the basis of the structural capacities and vulnerabilities displayed in terms of 6 dimensions. These dimensions record capacities for *self-awareness, self-regulation, defence, object awareness*, *communication*, and *attachment*. They are used to assess the patient's *level of integration* using the ratings "well-integrated", "moderately well-integrated", "poorly-integrated", and "disintegrated". The criteria for

these ratings are defined in the manual for all dimensions. In a final assessment, structure is given a global rating, on the same four-level basis. As Table 1 shows, each of the 6 dimensions has a number of subdimensions identifying the various aspects of the superordinate structural capacity in question. For example, the *capacity for self-regulation* dimension encompasses the subdimensions tolerance of affects, regulation of self-esteem, regulation of impulses, and anticipation (Rudolf, Oberbracht & Grande 1998).

3.2 Considerations bearing upon the use of the OPD

As noted above the usual devices for measuring change define improvements in the course of therapy in terms of the gradual alleviation of pathology (symptoms). In the case of the Operationalized Psychodynamic Diagnosis, however, this model has only limited significance. There are three reasons for this:

- **Studies on 12-week inpatient therapies (Grande, Rudolf & Oberbracht 2000) have shown that OPD findings are relatively stable over time and show very little change in relatively brief therapies. This is scarcely surprising if we remember that the OPD is designed to capture difficulties located at a deeper level in the patient's personality and hence less easy to change than symptoms or symptomatic behavior (cf. Schulte 1995; Grande & Jakobsen 1998).**
- **In findings of a psychodynamic nature the model of therapeutic "improvement" of pathological conditions has only limited relevance because here change does not take place in terms of "more“ or "less“ but rather along the lines of a qualitative reshaping or an enhanced integration of problematic aspects of the OPD profile. In the successful course of an analytic process a patient's central conflicts are not neutralized; it would be more accurate to say that they are constructively modified and better integrated in the important spheres of life. Nor does the central problematic relationship pattern become "diminished" in the course of a successful therapy; what happens instead is that it loses more and more of its compulsive character, involves less subjective suffering for the patient and is recast in qualitative terms.**
  With respect to the structural vulnerabilities of the type identified in the “Structure” axis of the OPD it would be more fitting to regard improvement in terms of the disappearance or reduction of pathological abnormalities. For example, in the course of successful treatment a patient's affect tolerance or self-object differentiation may change and this change may indeed be reflected in the rating scores. But here, too, it often seems clinically more appropriate, in the case of therapeutic success, to speak of an enhanced integration of certain

vulnerabilities, which by no means implies that the latter have simply vanished as a theme in the patient's life.

- **A further problem is that for individual patients not all the sections of the OPD profile may be equally relevant in change terms. There are always a small number of pathogenic conflicts and structural vulnerabilities that are especially significant for the specific problems of a given patient. Frequently, other abnormalities can be interpreted as secondary repercussions of these central problem areas. This small set of individually crucial areas can be regarded as nodal points in a network of dynamic interrelations on which other problems depend. They thus represent basal reference points for any treatment aiming at substantial therapeutic change.**

3.3 The focus concept

A technique for measuring change on the basis of the OPD must be designed in such a way as to take due conceptual account of the difficulties listed above. In our study we do this by defining *change as restructuring in the sense of a growing integration of specific problem areas that are of central significance for a patient's psychodynamics.* We assume that it is possible to define for every patient a limited number of such specific problems that can be used to observe therapeutic change. We also refer to these problem areas as “foci”, but it is essential to note that in the present context these are *research foci* and not therapeutic foci, in contrast to normal parlance.

The selection of these foci is undertaken via expert assessment by the external raters (cf. fig 1). The problem areas rated here as foci are those that are presumed to sustain both the patient's psychic/psychosomatic symptoms and his/her interpersonal problems. One problem area from the OPD spectrum is judged as being central and selected as a focus, in the sense that here something will have to change if the patient's problems are to be alleviated or dispelled. Defining the foci therefore is in the nature of establishing a case-related psychodynamic hypothesis specifying a patient's change-relevant characteristics.

Technically the procedure is that five central problems for each patient are selected from the OPD sectors Relationship, Conflicts, and Structure. This choice is based on the range of 30 potential problem areas listed in table 1. These areas derive from the core dysfunctional relationship pattern, the 8 conflicts and the 21 subdimensions from the Structure axis. In former studies (Grande, Rudolf & Oberbracht, 2000) we have established that the selection of 5 foci is sufficient to home in on the most important aspects of a patient's psychodynamic constitution. These studies have also demonstrated that in every case the habitual dysfunctional relationship pattern should be defined as one of the foci. The remaining problems are selected from the areas Conflict and Structure, with the proviso that at least one problem area be selected from each of these axes. Thus the selection of foci can be weighted in favour of conflicts or structural

vulnerabilities depending on the severity of the structural impairment displayed by a given patient. This reflects clinical experience of the way in which, depending on the nature and severity of an impairment, the diagnosis and treatment of patients will place greater emphasis either on structural features or on unconscious conflicts.

Fig. 2: The Heidelberg Structural Change Scale (HSCS)

Stages		Excerpt from the manual
1. Focus problem warded off	**exact** **1** **match** **1+** **tendency↓**	The problem is completely unconscious; experiences connected with it are evaded; problematic behavior is ego-syntonic; the patient has "no problems" with the critical area
2. Unwanted preoccupation with the focus	**tendency↑** **2-** **exact** **2** **match** **2+** **tendency↓**	Unpleasant feelings and thoughts in connection with the problem area can no longer be immediately rejected; but preoccupation with the problem is reluctant; external confrontations with the problem take place but are rejected as disturbances; no realisation that the problems might be associated with the patient's own person
3. Vague awareness of the focus	**Tendency ↑** **3-** **exact** **3** **match** **3+** **tendency↓**	Patient notices/suspects the existence of a problem that is part of him/herself and cannot simply be rejected; in the course of repetition the problem takes on a continuing existence; negative affects originate from the tension between the insistent nature of the problem and the patient's defensive/aversive attitude
4. Acceptance and exploration of the focus	**tendency↑** **4-** **exact** **4** **match** **4+** **tendency↓**	The problem starts to take on a new shape in the patient's consciousness; incipient indications of an active, "head-on" preoccupation with it; the problem can now be formulated as an "assignment" and hence be made the subject of therapeutic work; destructive, rejecting responses may interfere with this attitude but can no longer undermine it altogether

5. Deconstruction in the focus area	**tendency↑** **5-** **exact** **5** **match** **5+** **tendency↓**	Querying and disintegration of accustomed coping modes; uncertainty about evaluations of own person and others; perception of own limitations and deficiencies; resignation and moods of despair alternate with urges toward reparation; old modes are lost and cut off, new ones not yet accessible
6. ***Reorganization in the focus area***	**tendency↑** **6-** **exact** **6** **match** **6+** **tendency↓**	Abandonment and final relinquishing of accustomed coping modes; in his/her own experience patient is increasingly self-reliant and able to take in hand and assume responsibility for his/her own life in the problem area; increasingly conciliatory approach to problem area; problem solutions spontaneous and unexpected; re-integration
7. Integration of the focus problem	**tendency↑** **7-** **exact** **7** **match**	Dealing with the problem has become something natural; the area has lost its special significance in the eyes of the patient; the problem is something belonging to the past, preoccupying patient - as a memory

3.4 The Heidelberg Structural Change Scale (HSCS)

After selecting the foci the next step is to assess the state of therapeutic change the patient has reached with regard to these problem areas. For this purpose we use a modified form of the Assimilation of Problematic Experiences Scale (APES) by Stiles, Meshiot, Anderson & Sloan (1992). This scale enables us to describe more subtle changes in a patient's dealings with given structural problems. The term "assimilation“ here designates, with reference to Piaget, a process in which difficult experiences are acquired, integrated, and reshaped. The authors themselves conceptualize this process as being free of theoretical biases or allegiance to any specific therapeutic orientation. We have revised APES with an eye to more closely assimilating it to the exigencies of psychoanalytic treatment (Rudolf, Grande & Oberbracht 2000). The revision is in line with the logic set out in Freud's 1914 study "Remembering, Repeating and Working-Through“. Our modifications of APES are extensive, and hence we refer to this instrument as the *Heidelberg Structural Change Scale* (HSCS).

Each stage of the scale marks a therapeutically significant step, beginning with the increasing awareness of a problem area not perceived as such until then, extending through to the therapeutic working-through of the aspects and experiences associated with it, and from there to more basic changes resulting from it both in the patient's experience and in his/her concrete external behavior. With the aid of the scale patients are assessed as to the degree of structural change they have achieved at a given point in their treatment. A separate assessment is made on the scale for each of the five foci defined.

Figure. 3 gives an overview of the stages in the measurement of structural change. At the outset of therapy a semi-structured videotaped OPD interview (Janssen, Dahlbender, Freyberger et al. 1996) is used to rate the OPD, to select the five focal problems from the OPD focus list and to assess the patient’s ability to deal with these foci in terms of the Heidelberg Structural Change Scale (HSCS). At each new rating timepoint the 5 foci are re-assessed with respect to the HSCS on the basis of new interviews, thus pointing up the progress of restructuring within the separate focal areas. Two raters work independently at each stage of this rating process. After completion they are asked to arrive at a consensus rating on the basis of their independent judgments and to record this in a written commentary, thus producing a brief description of the level of therapeutic progress achieved for each of the focal areas. Fig. 9 and Table 3 give an example of the way the therapeutic changes for the 5 foci of an individual patient are recorded on the HSCS (for comments on this case see Section 4).

Fig. 4: Overview of Stages in Measurement of Structural Change

3.5 Reliability and validity

Investigations of the reliability of the method described here were undertaken in the framework of a study on inpatient therapy for patients with psychosomatic, neurotic and personality disorders (Rudolf, Grande, Oberbracht et al. 1996). We draw further material from the cases within the ongoing study reported on in this present article. The scores set out below are based on calculations collating data from both studies.

For the HSCS we have established interrater agreement of r=.77 (Pearson correlation) on the basis of N=306 individual focus ratings. – The reliability test for *focus selection* refers only to the 4 foci selected from the OPD sectors "Conflict" and "Structure" (axes III and IV). As the dysfunctional relationship pattern was pre-defined as a focus in all cases (see above) it is not possible to calculate rater agreement on it. On the basis of N=161 focus selections the result was a Kappa of .59. Given the large number of potentially selectable foci this figure is acceptable. It corresponds to a relatively conservative estimate of agreement because it takes no account of existing similarities between the foci *within* the 6 OPD Structure dimensions. As Table 1 shows, the subdimensions allocated to the Structure dimensions are relatively similar. A categorization of the foci on the basis of the 6 main dimensions produces an agreement result of a Kappa=.70.

Indications of the validity of the method can also be drawn from the inpatient study referred to. With reference to a study involving 49 patients, Grande, Rudolf, Oberbracht et al. (2000) report that the pre-post differences averaged for all 5 foci of a given patient show high correlation with the global outcome assessments of various members of the therapeutic team. As Table 2 shows, these correlations vary between .43 and .50. No other scale of change in this study had a higher degree of correlation with the global outcome. Indeed, the correlation between these assessments and the symptom change measurements "Somatization" (SOM), "Depression" (DEP), "Social Anxiety" (SA) and "Somatic Anxiety Symptoms" (SoA) was lower throughout (pre-post differences of the PSKB-Se, (Rudolf 1991b).

Tab. 2: The HSCS in relation to different assessments of therapeutic success

	*Global assessments of therapeutic success by...**		
Pre-post differences of... ⇓	Therapist	Therapeutic rounds	Staff
HSCS	.50**	.47**	.43*
Somatization	.49**	.34*	.35
Depression	.42**	.27	.26
Social Anxiety	.30	.11	.29

Somatic Anxiety Symptoms	.27	.25	.14

* N= max. 49 because of missings

These findings justify the conclusion that, taken in isolation, symptomatic changes do not suffice to capture what clinicians regard as therapeutic success. The study shows that the focus-related changes measured with HSCS map the global outcome of the treatment with greater clinical accuracy. This may have to do with the fact that clinicians (of a psychodynamic persuasion) do not rely on symptom improvement alone when assessing treatment success but inquire whether and to what extent patients have mode progress in the working-through of their central problem areas. Clinicians may have greater confidence in this form of therapeutic progress because it represents a sounder basis for further development after completion of therapy.

That this is indeed the case is borne out by the _ -year follow-up study by Grande, Rudolf, Oberbracht et al. (in preparation) on 39 patients from the same study. It showed that patients with a high score on the HSCS (averaged across 5 foci) at the end of therapy did better after discharge. The analysis of significant events in the external lives of the patient after inpatient treatment indicated that coping with those events was more adaptive if the patient had achieved improved HSCS scores in the course of therapy (highly significant correlation of .42). This can be interpreted as meaning that patients were better able to cope with the demands placed on them by their external lives if they had gained an awareness of the problematic tendencies interfering with their coping efforts in those areas. The study also demonstrates that this is not possible before stages 3 or 4 on the scale, when perception and/or acknowledgment of the focus problem has been achieved (cf. Fig. 2).

These findings show that the HSCS can be used reliably and maps changes at a good level of agreement with the assessment of therapy success by clinical experts. They also strongly suggest that improvements on the HSCS correlate with patients' increasing ability to regulate and cope with problem-related demands placed on them in their external lives. All in all, then, these studies supply evidence that the HSCS is an instrument which can be used to measure changes on a deeper level than the merely symptomatic. In the next section we take an individual instance of psychoanalytic psychotherapy to show how structural changes can be measured with this instrument.

4. A case study

4.1 The patient: clinical data

Mr. B, 22 years old, responded to his girlfriend's termination of their relationship with depression and suicidal leanings. He had been subdepressive for

a number of years before that and suffered from social anxieties impairing his career prospects. In the presence of others he regularly felt pressured and reacted with a markedly vegetative symptomatology. In compensation he cultivated notions in which he fantasized about being equipped with unusual intellectual abilities and artistic potential. After his parents' separation early in his life the patient grew up with his mother. His dependency on her was intensive and ambivalent. Despite their many conflicts he was hardly able to live a life of his own. He had occasional contact with his father, of whom he was critical but whom he also painfully missed.

According to ICD-10 this was classified as a depressive episode (F32.1) and a social phobia (F40.1). The following five problem areas from the OPD profile were established as foci for the patient: a.) the dysfunctional, maladaptive relationship pattern, b.) sexual/oedipal conflicts, c.) regulation of self-esteem, d.) contact, and e.) detachment (c. to e. from the Structure axis).

After separation, the patient embarked on a course of psychoanalytic psychotherapy involving 3 hours a week with a male analyst (recumbent). After 2.5 years and approx. 280 sessions, the therapy is now over and the data from it are available in their entirety. The follow-up studies have yet to ensue. The following delineates the course and the outcome of this treatment as reflected in the patient's self-assessment and the external examiners' ratings. The following scales and assessment instruments were used in the charting of the case:

1. Symptom changes mapped by three PSKB-Se Scales (Psychic and Social-Communicative Questionnaire, self-assessment version; Rudolf 1991b): “Somatic Anxiety Symptoms”, “Depression” and “Somatization”.
2. Inventory of Interpersonal Problems IIP (IIP-D German version - Horowitz, Strauß & Kordy 1993).
3. Heidelberg Structural Change Scale (HSCS)
4. Freely formulated clinical observations by the external raters on the selected "detaching" focus (e.) and the changes discernible in the course of therapy.

Naturally the documentation of this case as given here is not complete because a large number of additional data have not been taken into account. We refrain from alluding to the psychoanalyst's perspective and his assessment of the analytic process. We also neglect the health insurance data and what they tell us about the cost of the medical services provided and the number of days the patient was off work sick. Instead we limit our view to the assessments listed in

order to obtain a clearer picture of the ratio of symptomatic and structural changes displayed by Mr. B. In so doing we are also preparing the ground for Section 5 where we take a closer look at the relationship between psychotherapeutic and psychoanalytic treatment effects and refine the central hypothesis underlying this study.

4.2 Self-assessments of Mr. B.

Fig. 4 shows the progress of the symptomatology over the first two years of psychoanalysis, measured on the scales "Somatic Anxiety Symptoms", "Depression" and "Somatization" taken from the PSKB-Se (Rudolf, 1991b). The ratings are expressed as T-values; scores below 60 can be regarded as clinically “normal”. The figure shows a drop in the initially very high symptom scores to a normal range after a period between 6 months and 1 year; thereafter we have a reactivation of the symptomatology to a clinically abnormal score after 2 years. This applies notably to the scales “Depression” and “Somatic Anxiety Symptoms”. But after this temporary rise the scores on both scales ultimately return to the clinically normal range.

Fig. 4: Symptom change in the course of treatment

We find a very similar curve for the scales "over-introverted", "over-exploitable", "over-submissive", and for the IIP-D (Horowitz, Strauss & Kordy 1993), where the patient again had a quite high score at the outset. Here the reactivation trend in the second year of treatment was even more marked. But again we have a reduction of interpersonal problems toward the end of therapy.

Fig. 5: IIP scales "over-introverted", "over-exploitable" and "over-submissive”

Fig. 6: IIP summary scale

4.3 Comments and ratings by the external independent examiners

Figures 7 and 8 show the development of the patient from the perspective of the external raters assessing the changes in the 5 focus areas on the basis of videotaped OPD interviews and using the Heidelberg Structural Change Scale. The figures show the scores for the individual foci in the course of time. The scores for “relationship pattern”,“detachment” (Fig. 7), and “contact” (Fig. 8) show continuous improvement and by the end of treatment are in fact very good. The focus areas “oedipal-sexual conflicts” (Fig. 7) and “regulation of self-esteem” (Fig. 8) also show improvement but after one year remain constant at stage 4 of the Scale “acceptance and exploration of the focus”. Of note is the fact that after 1.5 and 2 years of therapy the patient had reached stage 5 “deconstruction in the focus area” first for one (Fig. 7: detaching, score 5-) and then for two (relationship and oedipal-sexual conflicts, scores 5- and 5) focal areas. The Manual (cf. abridged version in Fig. 2) speaks in this connection of uncertainty, resignation and despair, perception and ultimate acknowledgment of the patient's own limitations and impairments and depressive affects bound up with this.

Fig. 7: HSCS-rating for the foci “relationship”, “oedipal-sexual conflicts”, and “detaching”

Fig. 8: HSCS-rating for the foci “regulation of self-esteem” and “contact”

As mentioned earlier, after their HSCS rating the external examiners set down a detailed commentary in which the specific form taken by the focal areas and the focal changes are described with reference to the contents of the interviews. These commentaries have the advantage of providing a readily usable clinical impression of the patient and the therapy. Table 3 shows a condensed version of these commentaries for one of the focal problems of Mr. B. The focus in question is "detaching", figuring in the focus list (Fig. 3) as a structural capacity/vulnerability and allocated to the OPD structure dimension “attachment”. We see here a development in which over the period of 2.5 years the patient moves through a defensive/anxious and at the same time aggressive distancing from the object to a more intensive involvement accompanied by ambivalence and feelings of depression; subsequently a genuine detachment takes place in the aftermath of which the patient can draw on a sound degree of autonomy and for that reason can afford without risk to seek close relations with the objects and open out to them.

4.4 Synopsis of the case

We can summarize the course taken by our case example on the various observation levels as follows: in the first year the patient develops from a more reluctant engagement with his problem areas (Stage 2 of the Structural Change Scale, Fig. 2) to active acknowledgment and exploration of his problems (Stage 4). During this period the somatic, psychic and interpersonal complaints disappear almost totally. At the beginning of the second year of therapy we see a qualitative difference. The patient experiences his problems more intensively. On the basis of the more firmly established therapeutic relationship he now ventures to, as it were, expose himself to them. At the symptom level this goes hand and hand with a moderate reactivation of symptomatology and a clear increase in the interpersonal difficulties experienced by the patient. A phase of more intensive analytic working-through has begun, which we allocate to Stage 5 ("deconstruction in the focus area”) on the HSCS. Only after this phase of

instability do the changes set in that in clinical terms appear soundly "organic" and firmly rooted, thus promising to be more than temporary. The follow-up will show whether this interpretation is valid.

Tab. 3: External raters' commentaries on development in focal area "detaching" (condensed version)

	HSCS score	**Patient's management of focus problem**
Onset	2+	Despair, depression and thoughts of suicide following separation from girlfriend; inability to live a life of his own; responsibility for this attributed to mother
3 months	4-	First rift in the union with the mother through realization that she will die some time; acknowledgment of having a part in this dependence; purchase of own washing-machine; twin-like relationship with a friend
6 months	3-	Engagement with the focal problem has become vaguer again, much acting-out in this connection; first meeting with father; welcomes analysis break because of fear of dependence on analyst
1 year	4-	Reflection on his "fateful union" with the mother; hate and conciliatory reports on the mother, more differentiated, three-dimensional description of mother; interest in sexuality and own attractiveness; loosening of bonds with friend, more clearly delimited relationship; first emphasis of autonomy vis-à-vis interviewer
1.5 years	5-	Mourning over contradictory relationship with mother and guilt feelings; disappointed at relationship with father; feels wrongly treated in the analysis; makes indirect accusations in OPD interview (dryness of mouth!); expresses longing for a girlfriend

2 years	6-	Conciliatory moods where, though regretting what his mother was unable to give, he shows understanding for the restrictions she placed on him and takes a caring view of this; only very occasional defiant bids for delimitation; new self-assurance with regard to own abilities and potential; feels more autonomous
2.5 years	7-	Is able to empathize with mother and support her without worrying about self-delimitation; is also able to accept support from another woman; is in love again; feels enterprising and autonomous.

5. Structural and psychotherapeutic changes

In our discussion of a case example, stage 5 on the Heidelberg Structural Change Scale and the temporarily destabilizing processes bound up with it mark a caesura allowing a distinction between the effects of psychotherapies and psychoanalyses in terms of ideal types. Clinical experience shows that in psychoanalyses external changes may take a long time to materialize, appearing spontaneously at a later date when a solution has had sufficient time to mature inwardly. In terms of the process model we have developed, this occurs at Stages 6 and 7 when old, consolidated defence or coping structures have disintegrated (Stage 5).

The spontaneous character of changes deriving from an analytic process is quite conceivably one feature clearly distinguishing psychoanalyses from other forms of psychotherapeutic treatment aiming at a specific therapeutic change either via focussing or direct educative intervention. Fig. 9 shows how various modes of change can be allocated to certain sections of the HSCS.

Fig. 9: Modes of therapeutic change

- *Educative mode:* At Stages 1 and 2 of the Scale, positive changes are most likely to be achieved if the therapeutic approach is symptom-related and

educative. Here the patient is more extrinsically motivated in his attempts to bring about change and (in productive instances) behaves in a compliant, cooperative way. Correspondingly the changes themselves are symptom-related. At a secondary level, however, the decrease of symptomatic impairment can definitely trigger notable other effects, e.g. via a gain in the subjective awareness of personal competence and an improvement in self-esteem, which in its turn may then generate further favorable effects of a nature not necessarily specific to the intentions of the therapy.

- *Coping mode:* Distinct from this is a coping mode where the therapeutic approach is geared to the inner psychic problem domain connected with the symptomatology and sets out to uncover this focally. The patient's insight into his/her own problem enables him/her to attain greater ability to manage their own problems in the form of conscious regulation and to bring about change via intentional endeavor. Here the motivation is more intrinsic. From this mode broader positive effects may derive, as described above (coping mode).
- *Structural change mode:* The third mode is the mode of structural change proper. Here the therapeutic approach is basically open and characterized by the willingness of the therapist to take comprehensive account of the personality of the patient in its conscious/unconscious forms of expression and to allow for an analytic process. Here again, the patient's changes are intrinsically motivated but essentially they are not consciously desired but transpire spontaneously, sometimes surprising the patients themselves. The effects triggered by this mode are primarily broad in scope and at the same time specific to the equally broad therapeutic intention.

It needs to be stressed that these modes can by no means be paired off with therapy techniques in a clear-cut manner. Within behavioral therapy, for example, there are techniques that would certainly qualify for inclusion in the coping mode. Equally, certain intrinsically psychodynamic therapies may display educative elements or in certain cases initiate an analytic process, even though the setting does not comply with the classical requirements. For this reason the change modes in the Figure are allocated to the stages on the Scale in a way intended to be approximate and overlapping.

In principle, however, our working hypothesis for the study as a whole is *that in psychoanalysis there is a higher incidence of developments corresponding to the third change mode (structural change) and described by Stages 5 to 7 on the State of Structural Change Scale*. Vice versa, for psychotherapies we expect a higher frequency of changes taking place at Stages 3 and 4 and corresponding to

the coping mode. This specific hypothesis is the central distinctive assumption of the study and will be subjected to statistical verification via comparison between the groups.

We anticipate that this approach will do more than merely furnish global evidence for the superiority of one therapy form over another. We expect it to supply a better understanding of the way in which various different processes achieve their effects, of the processes they set in train, the likelihood not only of a good sustainable therapy outcome but also of the risks of standstill and failure. The structural change model developed by the authors and condensed in the Scale represents a viable method of imaging change processes beyond the symptom level and hence of capturing the specific outcomes traceable to psychoanalysis in the therapeutic change process.

References

Arbeitsgruppe OPD (Ed.) (1998): Operationalisierte Psychodynamische Diagnostik. Grundlagen und Manual. Bern etc.: Huber.

Becker, P. (1989): Der Trierer Persönlichkeitsfragebogen TPF. Göttingen, Hogrefe.

Benjamin, L.S. (1974): Structural analysis of social behavior. Psychological Review 81, 395-425

Derogatis, L.R., Lipman, R.S., Covi, L. (1975): SCL 90. An outpatient psychiatric rating scale. Psychopharmacology Bulletin 9, 13-28. (The revised version, SCL 90-R, by L.R. Derogatis appeared in 1983 in Towson, MD as an academic print.)

Dilling, H., Mombour, W., Schmidt, M.H. (1991): Internationale Klassifikation psychischer Störungen. Bern etc.: Huber.

Freud, S.F. (1914): Remembering, Repeating and Working-Through. (Further Recommendations on the Technique of Psycho-Analysis, II). Standard Ed., 12, 147.

Grande, T., Porsch, U., Rudolf, G. (1988): Muster therapeutischer Zusammenarbeit und ihre Beziehung zum Therapieergebnis. Zeitschrift für Psychosomatische Medizin und Psychoanalyse 34, 76-100.

Grande, T., Burgmeier-Lohse, M., Cierpka, M., Dahlbender, R.W., Davies-Osterkamp, S., Frevert, G., Joraschky, P., Oberbracht, C., Schauenburg, H., Strack, M., Strauß, B. (1997): Die Beziehungsachse der Operationalisierten Psychodynamischen Diagnostik (OPD) - Konzept und klinische Anwendungen. Zeitschrift für psychosomatische Medizin und Psychoanalyse 43, 280-296.

Grande, T., Rudolf, G., Oberbracht, C. (1997): Die Praxisstudie Analytische Langzeittherapie. Ein Projekt zur prospektiven Untersuchung struktureller Veränderungen in Psychoanalysen. In: Leuzinger-Bohleber, M., Stuhr, U. (Eds.): Psychoanalysen im Rückblick: Methoden, Ergebnisse und Perspektiven der neueren Katamneseforschung. Gießen: Psychosozial Verlag, 415-431.

Grande, T., Jakobsen, Th. (1998): Zur Notwendigkeit einer psychodynamischen Diagnostik und Veränderungsmessung in quantitativen Studien zur analytischen Psychotherapie und Psychoanalyse. In: Fäh, M., Fischer, G. (Eds.): Sinn und Unsinn in der Psychotherapieforschung - Eine kritische Auseinandersetzung mit Aussagen und Forschungsmethoden. Gießen: Psychosozial-Verlag, 125-137.

Grande, T., Rudolf, G., Oberbracht, C. (2000): Veränderungsmessung auf OPD-Basis -Schwierigkeiten und ein neues Konzept. In: Schneider, W., Freyberger, H-J. (Eds.) Was leistet die OPD. Bern etc.: Huber, 148-161.

Grande, T., Rudolf, G., Oberbracht, C., Jakobsen, T. (2000): Therapeutische Veränderungen jenseits der Symptomatik – Wirkungen stationärer Psychotherapie im Licht der Heidelberger Umstrukturierungsskala. Zeitschrift für Psychosomatische Medizin und Psychotherapie. Zeitschrift für Psychosomatische Medizin und Psychotherapie (in press).

Grande, T., Rudolf, G., Oberbracht, C., Pauli-Magnus, C. (in preparation): Progressive Veränderungen im Leben des Patienten nach Entlassung – durch welche Effekte stationärer Psychotherapie werden sie gefördert?

Grawe, K., Donati, R. & Bernauer, F. (1994): Psychotherapie im Wandel. Von der Konfession zur Profession. Göttingen, Hogrefe.

Horowitz, L.M., Strauß, B., Kordy, H. (1993): Inventar zur Erfassung interpersoneller Probleme. Manual der deutschen Version.

Howard, K.I., Kopta, S.M., Krause, M.S., Orlinsky, D.E. (1986): The dose-effect relationship in psychotherapy. American Psychologist, 41, 159-164.

Janssen, P.L., Dahlbender, R.W., Freyberger, H.J., Heuft, G., Mans, E.J., Rudolf, G., Schneider, W., Seidler, G.H. (1996): Leitfaden zur psychodynamisch-diagnostischen Untersuchung. Psychotherapeut 41, 297-304.

Keller, W., Dilg, R., Westhoff, G., Rohner, R., Studt, H.H. (1997): Zur Wirksamkeit ambulanter jungianischer Psychoanalysen und Psychotherapien – eine katamnestische Studie. In: Leuzinger-Bohleber, M., Stuhr, U. (Eds.): Psychoanalysen im Rückblick: Methoden, Ergebnisse und Perspektiven der neueren Katamneseforschung. Gießen: Psychosozial Verlag, 432-453.

Rudolf, G. (1991a): Die therapeutische Arbeitsbeziehung. Untersuchungen zum Zustandekommen, Verlauf und Ergebnis analytischer Psychotherapien. Unter Mitarbeit von T. Grande und U. Porsch. Berlin etc.: Springer.

Rudolf, G. (1991b): PSKB-Se - ein psychoanalytisch fundiertes Instrument zur Patienten-Selbsteinschätzung. Zeitschrift für Psychosomatische Medizin und Psychoanalyse 37, 350-360

Rudolf, G.; Grande, T. (1997): Praxisstudie Analytische Langzeittherapie. Forschungsprojekt zur Untersuchung der Effektivität und Effizienz höherfrequenter analytischer Langzeitpsychotherapien. Projektantrag.

Rudolf, G., Grande, T., Oberbracht, C., Jakobsen, T. (1996): Erste empirische Untersuchungen zu einem neuen diagnostischen System: Die Operationalisierte Psychodynamische Diagnostik (OPD). Zeitschrift für Psychosomatische Medizin und Psychoanalyse 42, 343-357.

Rudolf, G., Laszig, P., Henningsen, P. (1997): Dokumentation im Dienste von klinischer Forschung und Qualitätssicherung. Psychotherapeut 42: 145-155.

Rudolf, G., Oberbracht, C., Grande, T. (1998): Die Struktur-Checkliste. Ein anwenderfreundliches Hilfsmittel für die Strukturdiagnostik nach OPD. In: Schauenburg, H., Freyberger, H.J., Cierpka, M., Buchheim, P. (Eds.): OPD in der Praxis. Konzepte, Anwendungen, Ergebnisse der Operationalisierten Psychodynamischen Diagnostik. Bern etc.: Huber, 167-181.

Rudolf, G.; Grande, T.; Oberbracht, C. (2000): Die Heidelberger Umstrukturierungsskala. Ein Modell der Veränderung in psychoanalytischen Therapien und seine Operationalisierung in einer Schätzskala. Psychotherapeut 45. 237-246.

Sandell, R., Blomberg, J., Lazar, A. (1999): Wiederholte Langzeitkatamnesen von Langzeitpsychotherapien und Psychoanalysen. Zeitschrift für Psychosomatische Medizin und Psychotherapie 45, 43-56.

Schepank, H. (1995): BSS - Der Beeinträchtigungs-Schwere-Score. Göttingen: Beltz-Test.

Schulte, D. (1995): How treatment success could be assessed. Psychotherapy Research 5 (4), 281-298.

Seligman, M.E.P. (1995). The effectiveness of psychotherapy: The Consumer Reports study. American Psychologist, 50, 965-974.

Stiles, W.B., Meshot, C.M., Anderson, T.M., Sloan, W.W. (1992): Assimilation of problematic experiences: The case of John Jones. Psychotherapy Research 2, 81-101.

Strupp, H.H.; Schacht, T.E.; Henry, W.P. (1988): Problem-Treatment-Outcome Congruence: A principle whose time has come. In: Dahl, H., Kächele, H., Thomä, H. (Eds.) Psychoanalytic Process Research Strategies. Berlin etc.: Springer-Verlag, 1-14.

Tress, W. (Hg.) (1993): SASB- Die Strukturale Analyse Sozialen Verhaltens. Heidelberg: Asanger.

Contribution to the Measurement of Mode-Specific Effects in Long-Term Psychoanalytic Psychotherapy

Dorothea Huber and Guenther Klug

In psychotherapy research generally agreed upon standards have been developed for the assessment of outcome during the last few years. According to Lambert and Hill (1994) the state of the art requires a) assessment from different sources (self-reporting, trained observers, relevant others, therapist rating, institutional); b) application of different, non-reactive technologies (global ratings, specific symptom index, observer ratings, physiological measures, life records); c) atheoretic, pragmatic measures; d) bi-directional measures; e) multidimensional measures (intrapsychic, interpersonal, psychosocial); and f) follow-up measurements. Strupp, Horowitz and Lambert (1997) developed criteria for a core battery to measure outcome: a) clear and standardised procedures for administering and scoring the instrument; b) norms for patient and non-patient populations; c) demonstrated reliability; d) demonstrated validity; e) demonstrated sensitivity to change; f) demonstrated feasibility in clinical settings; g) systematic rater-training and calibration tapes available if administered by trained clinicians; h) atheoretic measures; i) multimodal measures; j) categorical and dimensional measures; and k) measurement before, during and after treatment.

Both approaches have to cope with the problem of grasping the mode-specific and therefore theory-bound effects of one specific therapeutic modality as well as applying theory-free, pragmatic measures not confined to the perspectives of one specific theory of change.

Schulte (1995) tried to elucidate the dilemma by distinguishing four levels within the general concept of disease: a) causes (biological, psychological, sociological); b) disease (pathological changes in the person); c) illness (symptoms, complaints, medical findings); and d) consequences (sick-role and impairment of normal role behaviour). He argues that on level a) causes and level b) disease only school-specific, hence theory-bound measures are

appropriate, while on level c) illness and level d) consequences theory-free measures are appropriate.

Reviewing the literature on the outcome of psychoanalysis and psychoanalytic psychotherapy, all authors (e.g. Roth and Fonagy, 1996; Vaughan, Marshall, MacKinnon, Vaughan, Mellman and Roose, 2000) agree that, although some evidence of the effectiveness of these treatments has been accumulated, there still is a regrettable lack of studies that meet the requirements of modern empirical research.

One of their main points of critique has been the absence of appropriate measures to encompass the more ambitious aims of psychoanalysis and psychoanalytic psychotherapy. The outcome studies of the 70¥s and 80¥s mostly applied global assessments of therapeutic benefit that were not able to capture the specific effects of psychoanalysis and psychoanalytic psychotherapy. The outcome measures, mostly expert ratings of interviews with the patient and the treating analyst, did not meet modern research standards such as, for example, reliability, validity and sensitivity to change and could only cover dimensions of change in a very restricted way only, using vague categories. Bachrach, Galatzer-Levy, Skolnikoff and Waldron (1991) therefore concluded in their critical survey of outcome studies: ì...the research methods, especially of the clinical-quantitative studies, reflect the state-of-the-art of the 1950¥s and 1960¥s more than currently available methods.î Very much in the same vein is Luborsky, Diguer, Luborsky, Singer, Dickter and Schmidt¥s (1993) critique of unsuitable outcome measures because they ìdo not make an adequate distinction between short-term and long-lasting improvement, nor do they make a distinction between the parallel related changes referred to as non-structural and structural changeî.

Another decisive point is that psychotherapy research shifted from the investigation of the outcome of psychotherapy to the connection between process and outcome. This brings about a new challenge for the conceptualisation of outcome measures because the link between outcome and process has to be observed attentively. As a principle that tries to meet the needs of process-outcome research, Strupp, Schacht and Henry (1988) proposed the ìProblem-Treatment-Outcome Congruenceî (P-T-O Congruence method). Dahl (1988) elaborated on it thus: ìthis principle (says) that the description and representation, theoretically and operationally, of a patientís conflicts, of the patientís treatment, and of the assessment of the outcome, must be congruent, which is to say, must be represented in comparable, if not identical terms.î Therefore an outcome

measure is needed that both grasps the specific effects of psychoanalysis and psychoanalytic psychotherapy, and is able to link the outcome, on a conceptual basis of a commonly agreed upon theory of change to the process, making ìpsychoanalytic changeî more conceivable that way.

Psychoanalysis and psychoanalytic psychotherapy used to understand Schulte¥s level b) disease (pathological changes in the person) as ìpsychic structureî and coined the term ìstructural changeî for changes on that level. Pulver (1991) defined psychic structure ìas any organisation of mental contents and processes which, in a systematic way, carried out the various tasks of the psycheî. According to Piaget (1970) its basic principles are wholeness, transformation and self-regulation. The ìOperationalized Psychodynamic Diagnosticsî group (OPD-Task Force, 2001) characterised ìpsychic structureî by six features: a) capacity for self-reflection; b) capacity for self-management; c) capacity for defense; d) capacity for object perception; e) capacity for communication and f) capacity for binding.

Structural change as ìthe generally accepted goal of psychoanalysisî (Moore & Fine, 1990) and also to some degree of psychoanalytic psychotherapy (e.g. Kernberg, 1991; Wallerstein, 1986), is an explanatory construct that tries to capture the specific effects of psychoanalysis and psychoanalytic psychotherapy relating them to the concept of psychic structure and their modifications by psychoanalytic treatment. Structural change signifies a type of change, beyond symptoms and manifest behaviour, rooted in the matrix of both. Kernberg (1991) defines it as a ìsignificant modification in the unconscious intrapsychic conflicts underlying symptom formation. Change in the underlying unconscious intrapsychic structures is usually revealed in shifts in the equilibrium of ego, superego, and id, with a significant expansion of the system ego and a corresponding reduction of the pressures of the unconscious superego and id î.

The application of such global ego-psychological concepts in follow-up studies revealed, however, that they had to be more clearly and differentially operationalised and had to be assessed by experts in order to grasp the more subtle effects of psychoanalytic treatments (Wallerstein, 1986). As to the technology (Lambert & Hill, 1994) of the instrument and above all its non-reactivity we would like to point out that only expert clinicians will be able to assess adequately the transference- related cognitive distortions that regularly take place in follow-up interviews (Pfeffer, 1959).

In summary we conclude that an empirically-founded outcome measure that intends to grasp changes beyond symptoms and behaviour in a process-outcome context, must a) be based on expert judgement; b) be able to grasp different functions of psychic structure and their modifications; c) be able to rule out the influence of symptoms on psychic functioning; d) be able to assess the influence of transference-related cognitive distortions in the follow-up interviews; e) be based on concepts agreed upon by adherents of different psychoanalytic schools; f) have satisfactory psychometric qualities (reliability, validity, sensitivity to change, clinical significance), and g) must meet the requirements of the P-T-O principle.
We consider the Scales of Psychological Capacities, SPC (Wallerstein, 1991; Wallerstein, unpublished manuscript) as the measure that complies best with these standards and that has sufficiently proven its psychometric qualities (DeWitt, Hartley, Rosenberg, Zilberg, & Wallerstein, 1991; Dewitt, Milbrath, & Wallerstein, 1999).

Therefore we decided to apply the Scales of Psychological Capacities (SPC) to a comparative process-outcome study of psychoanalysis and psychoanalytic psychotherapies, the Munich Psychotherapy Study (MPS), which we present next.

The Munich Psychotherapy Study, MPS (Huber, Klug, & von Rad, 1997; Huber, & Klug, 1999; Huber, Klug, & von Rad, 2000) is a process-outcome comparison study in progress that tries to answer two questions:-

1. Are there any differences in effectiveness between psychoanalysis and psychodynamic psychotherapy? And if so: are those changes brought about by psychoanalysis based on ìstructural changeî and as a result of this, are they more profound and more stable than those brought about by psychodynamic psychotherapy ?

2. Are there any links between therapeutic process and outcome? And if so: what are they?
In order to answer the first research question a randomised controlled design was chosen to compare the two experimental groups:

1. A group of patients treated with psychoanalysis (PA) taking place 3 times a week in a recumbent position with an average duration of 240 - 300 hours.

2. A group of patients treated with psychodynamic psychotherapy (PT) place once a week in a seated position with an average duration of 80 - 120 hours.

As already stated the effectiveness of the two treatments can only be evaluated correctly if the patients are assigned at random to the two experimental groups. Because of the relatively small number of patients in each group (N=30) a strict allocation at random would lead to an uneven distribution of important patient variables, which was one of the main issues of the NIMH depression study (Elkin et al., 1989); we therefore decided to stratify the patients with regard to severity of symptoms and age. We set high value on the therapies being assigned at random and not the therapists so as not to interfere with the important, individual patient-therapist match.

Each patient presenting at the outpatient department of the Institute of Psychosomatic Medicine, Psychotherapy and Medical Psychology of the Technical University of Munich who met the inclusion criteria received an extensive clinical intake interview that was audio-recorded. Based on this recorded interview a board of 3 experienced psychoanalysts (so called ìindication boardî) decided if the patient could be randomly assigned to one of the two experimental groups. This decision process was documented as precisely as possible.

The inclusion criteria were as follows: between 25 and 45 years of age; ICD-10 diagnosis of depressive episode or recurrent depressive disorder; Beck-Depression-Inventory, BDI (Beck, Ward, Mendelson, Mock, & Erbaugh, 1961) > 16; previous psychotherapy to be finished at least 2 years before entering the study; not on anti-depressant medication; living in Munich or nearby; adequate German language skills.

The 10 therapists who participate in the study are experienced psychoanalysts and psychotherapists in private practice and have been working with patients for at least 5 years. They were trained at an approved institute and graduated there. They only apply therapies in which they are experienced and are not obliged to apply a therapeutic modality he or she does not consider as suitable for a particular patient.

The data come from three different sources of observation: the patient, the therapist and the researcher (ìexternal investigatorî). The test battery of outcome measures is adapted from the core battery suggested by the Society for

Psychotherapy Research (SPR), published by Grawe and Braun (1994) and chosen to be comparable with other ongoing studies.

As already stated a major concern of the study is to measure not only symptoms and behaviour, but also mode-specific effects and therefore, special instruments to measure structural change and individual therapeutic goals are administered. Structural change is measured with the Scales of Psychological Capacities (SPC); individual goals are assessed by means of the Goal Attainment Scaling method, developed by Kiresuk and Sherman in 1968, which in the Heidelberg-Study of von Rad, Senf and Br‰utigam (1998) yielded an interesting discrimination between psychoanalysis and psychotherapy.

The procedural plan (schedule) of the study is shown in table 1:

Table 1

Procedural Plan of the MPS (see text for abbreviations of the instruments)

Pre-treatment Measurement External investigator 1 and patient: intake-interview, ICD-10 and DSM-IV diagnosis, GAF, BADO, BDI (>16), BSS, HAMD Board of three experienced analysts: decision on patientís inclusion in the study. External investigator 1 and patient: SPC-interview; informed consent

Patient: self-rating questionnaires: BDI, SCL-90-R, IIP, FKBS, INTREX, SOZU, BADO, FLZ, FPI-R. External investigator 1 and patient: assessment of individual goals (goal attainment scaling, GAS)

Referral to therapist: Therapist: documentation of diagnosis, psychodynamic hypothesis, level of personality organisation, treatment goals, prognosis, HAQ-T Process Measurement, Audio-recording of every session

Patient: self-rating questionnaires: BDI, SCL-90-R, IIP, GAS and HAQ-P every 6 months

Therapist: therapy accompanying card to be filled out after every session; periodical process rating scale with HAQ-T every 6 months Post-treatment

Measurement: External investigator 2 (ìblindî for applied therapy) and patient: post-treatment interview, SPC-interview, life-events checklist, ICD-10 and DSM-IV diagnosis, GAF, BSS, HAMD, BADO

Patient: self-rating questionnaires: BDI, SCL-90-R, IIP, FKBS, INTREX, SOZU, BADO, FLZ, FPI-R, GAS, VEV, HAQ-P

Therapist: periodical process rating scale and HAQ-T, assessment of termination of treatment. Follow-up Measurement

(annually) External investigator 2 and patient: follow-up interview, SPC-interview, life-events checklist, ICD-10 diagnosis, BSS, GAF, HAMD, BADO

Patient: self-rating questionnaires: BDI, SCL-90-R, IIP, FKBS, INTREX, SOZU, BADO, FLZ, FPI-R, GAS, VEV

At the end of the intake-interview with ICD-10 and DSM-IV diagnosis the external investigator completes the Global Assessment of Functioning Scale (GAF, DSM-IV axis 5; American Psychiatric Association, 1994), the Symptom Severity Score (BSS; Schepank, 1995), the Hamilton Rating Scale for Depression (HRSD, Hamilton, 1960) and the Basic Documentation of the German College of Psychosomatic Medicine (BADO, our version described by Huber, Henrich, & von Rad, 2000), including the rating of the psychic structure of the patient (axis 4: Structure of the Operationalized Psychodynamic Diagnostics, OPD; OPD-Task Force, 2001). After a positive decision by the ìindication boardî and the obtaining of ìinformed consentî by the patient the external investigator interviews the patient with a semi-structured SPC-interview to obtain the appropriate information to score the SPC-scales. In the third pre-treatment session the external investigator and the patient assess together the individual goals the patient wishes to achieve during the therapy. The patient is assigned to one of the experimental groups after this intake procedure, so that the external investigator is ìblindî to therapeutic modality during the pre-treatment assessment.

Before the treatment starts the patient completes the following self-rating questionnaires: Symptom-Check-List (SCL-90-R, Derogatis, Lipman, & Covi, 1975), Beck-Depression-Inventory (BDI, Beck et al., 1961), Inventory of Interpersonal Problems, short version (IIP-C, Horowitz, Rosenberg, Bauer, Ureno, & Villasenor, 1988), Introject questionnaire (INTREX, Benjamin, 1974), Questionnaire for Coping Strategies (FKBS, Hentschel, 1998), Freiburg Personality Inventory, revised version (FPI-R, Fahrenberg, Hampel, & Selg, 1989), Life Satisfaction Questionnaire (FLZ, Huber, Henrich, & Herschbach, 1988), Basic Documentation of the German College of Psychosomatic Medicine (BADO, our version described by Huber, et al., 2000), Social Support Questionnaire, short version (F-SOZU-K-22, Sommer & Fydrich, 1991).
The therapist completes the Helping Alliance Questionnaire (HAQ-T Alexander & Luborsky, 1986) and a documentation form with psychodynamic diagnoses, main defences, level of personality organisation, motivation, main psychodynamic hypotheses, treatment goals and prognosis.
During the ongoing therapeutic process neither the patient nor the therapist is contacted personally so as not to interfere with the process too much; of course research itself inevitably influences the process. The process measures are sent to

the patient and psychotherapist by mail every 6 month. Each therapy session is audio-recorded.

Measurement points for the outcome measures are at pre-treatment, at post-treatment and at follow-up each year after end of treatment. The external investigator 2 at post-treatment and follow-up is not the same as at pre-treatment and is ìblindî to the therapeutic modality ñ although the possibility of remaining ìblindî during a clinical interview has been questioned in the literature (Luborsky et al., 1999). At post-treatment and follow-up the pre-treatment instruments are applied again, along with a retrospective life-event checklist and a self-rating questionnaire: Change in Experiencing and Behaviour, VEV (Zielke & Kopf-Mehnert, 1978).

There follows a more detailed description of the Scales of Psychological Capacities (SPC), the measure of mode-specific effects of psychoanalysis and psychoanalytic psychotherapy (ìstructural changeî) applied in our study.

The Scales of Psychological Capacities, SPC are an expert-rating measure that, whilst theoretically-informed but not theory-specific, evaluates the level of psychic structure. They have been developed from the research methodology of the Psychotherapy Research Project (PRP) of the Menninger Foundation (Wallerstein, 1986) and try to operationalise the concept ìpsychic structureî and ìstructural changeî as independently as possible of the differing theoretical perspectives in psychoanalysis so that they are able to assess reliably the specific changes after psychoanalyses and psychoanalytic psychotherapies. Bound to an empirical research strategy these psychological capacities are designed to be as low-level (experience-near) constructs as possible, and readily inferable from observable behaviours and conscious states of mind so that underlying intrapsychic structures and their changes after treatment can be reliably captured.

The Psychological Capacities Scales consist of 17 dimensions, 14 of which are divided into 2 subdimensions and 2 of which into 3 subdimensions; 1 dimension is not divided. The assessment of all 36 subdimensions is based on a tape-recorded, one-hour clinical intake interview together with a one- to two-hour semi-structured SPC-interview with probe questions, developed by the test author and his group. The material gained this way is scored for each subdimension on a 7-point scale from 0 for ìnormalî or fully adaptive functioning to 3 for functioning seriously and obviously disturbed, with half points in between. The dimensions are constructed one subdimension to be

designated for different degrees of inhibited and another for different degrees of exaggerated functioning. Both directions have to be assessed and both subdimensions can be scored simultaneously. The rating procedure requires an extensive manual with a detailed description of each subdimension together with one or more clinical vignettes to anchor each scale point.

Although inter-rater reliability (DeWitt et al., 1999), content validity (DeWitt et al., 1991), and convergent validity (DeWitt et al., 1999) of the SPC have already been examined there is a lack of studies that prove their feasibility for German research projects as well as a lack of discriminant validity studies.

METHOD

Any outcome measure needs to have sufficient evidence that reliability and validity are warranted as the basic psychometric properties. Developers of outcome measures must strike a balance between these two psychometric qualities to be able to offer instruments to researchers that meet their demands.
Comparing the SPC with instruments that measure interpersonal functioning and personality structure can evaluate convergent validity. Discriminant validity is evaluated in the present study by comparing the SPC with instruments measuring symptomatology.

We expect a zero or only weak correlation between the Scales of Psychological Capacities (SPC) and the construct-distant measures and a moderate but not very high correlation between the Scales of Psychological Capacities (SPC) and the construct-near measures.

Validity study I

As this pre-study has already been published (Huber & Klug, 1997; Huber, Klug, & von Rad, 2000), we provide only a short description of the procedure. The sample consisted of a homogenous group of 41 depressed patients of the MPS study (see above). Two judges (DH & GK) rated the 41 SPC-interviews from the audiotapes. Both are psychoanalysts with a completed analytic training and many years of professional experience. They trained themselves with the SPC-manual and rated the first three interviews together. Afterwards they had recalibration sessions after every fifth rating. The construct-distant instruments were the SCL-90-R, a self-rating symptom inventory constructed to assess the psychological and symptom status of psychiatric patients on 9 scales and a global severity index and the BSS, and the GAF; the BSS is an observer rating scale that evaluates the

impact of psychic illness on three dimensions: physical, psychological and social. The GAF (recent and highest level of functioning in the last year) is rated by the external investigator as well and is an internationally used scale.

As a construct-near measure we used the Inventory of Interpersonal Problems (IIP) in its self-rating form, and additionally in an observer-rating form (Horowitz, personal communication) in order not to contaminate the results by changing the source of observation for the only measure used. The IIP is an internationally established instrument for assessment of interpersonal problems and concerns.

Inter-rater reliability study and validity study II

We performed a second construct validity study of the SPC as a replication of the above presented convergent and discriminant validity studies to further examine validity three years later. This study consisted of another sample of the MPS, a different interviewer, different raters who were trained according to the formal method (Mercer & Loesch, 1979), and of additional construct-near and construct-distant instruments. As the validity of any diagnostic instrument presupposes a reasonable degree of inter-rater reliability, we started this time with an inter-rater reliability study. Both studies are not yet published and therefore will be presented in more detail here.

The sample consisted of a homogenous group of 47 depressed patients, between 25 and 45 years. The diagnosis was made by an experienced clinician (psychiatrist and psychotherapist) using the IDCL classification schema for ICD 10 and DSM IV diagnosis (Hiller, Zaudig, & Mombour, 1995) after having discussed the cases with two other experts. According to the intake criteria all patients received a diagnosis of a depressive disorder. A description of the socio-demographic data and the ICD 10 diagnosis is given in table 2.

Table 2
Description of the sample (N = 47)

age (mean, sd) 34,6 (6,5)
Romantic relationship 60%
gender children

employment
  full time 63%
  part time 11%
  unemployed 4%

other 22%

female 66%
male 34%
marital status
single 66%
married 19%
divorced 15%

1st diagnosis (ICD-10)
F 32.1 depressive episode ñ moderate 43%
F 32.2 depressive episode ñ severe 17%
F 33.1 recurrent depressive disorder ñ moderate 23%
F 33.2 recurrent depressive disorder - severe 17%

2nd diagnosis (ICD-10)
F 34.1 dysthymia 47%
other 8%
no 45%

The 38 patients in the inter-rater reliability study were a subsample of the above described sample.

Procedure: The authors of this paper have attended a rater-training according to the formal method (Mercer & Loesch, 1979) with the PRP-II group in San Francisco and have afterwards trained three German raters. After every 5th patient there was a recalibration session where the three judges and the two trainers met to correct for judges¥ drift.

The two raters, who came out with the best reliability scores, rated the scales for the validity study II. If they disagreed more than one scale-point, a senior rater (one of the authors) rated the scale again ñ a method recommended by Jones, Cumming and Horowitz (1988).

Instruments: An instrument already used in the pre-study (validity study I) for construct-validity of the SPC, the IIP, short form (IIP-C) was used again. Additionally to the pre-study we compared the SPC with the FKBS measuring 5 defense mechanisms / coping strategies. A personality questionnaire (FPI-R) with 12 scales was added.

Besides these self-rating questionnaires, the external investigator rates the psychic structure of the patient (OPD, axis 4) on a 4-point scale (good, moderate, low integrated, disintegrated).

We again assessed discriminant validity with the SCL-90-R. Additionally the Beck Depression Inventory (BDI) was used for assessment of the severity of depression.

Statistics: The inter-rater reliability between the three raters was calculated by means of Intra Class Correlation Coefficient (ICC, Shrout & Fleiss, 1979) for all subdimensions separately.
Correlations between the SPC and the tests for discriminant and convergent validity were calculated with the Pearson Correlation Coefficient. Because of the large amount of correlations computed, we decided to interpret only findings on a 0.1% level of significance.

RESULTS

Validity study I: At first content validity was examined by prospectively (that means before the respective subdimensions were scored on empirical material) assessing, from a clinical point of view, those subdimensions which would be expected to be most highly scored. This clinically expected ìprototypicî profile of depressive patients was compared with the empirically found profile. With one exception, the empirically found mean profile of depressive patients was correctly predicted by the clinicians, very clearly demonstrating a high consensus between clinical judgement and SPC ratings.

The results of the discriminative and convergent construct validity study showed that there were no significant correlations between the SPC and symptoms measured by the SCL-90-R, the BSS and the GAF. There were 10 medium correlations (rr: .49- .64) between the SPC and the IIP scales, when IIP was rated by the patient (self-rating). There were 20 significant (?) correlations between SPC and IIP scales when IIP was rated by the external investigator (judge rating). This can be readily understood, as it is well known, that there is a higher correlation between data from the same source of observation than between the content of the scales. These results confirm the hypothesis that the SPC measure is relatively independent of current symptoms and measures something similar but not identical to interpersonal functioning (Huber, Klug, & von Rad, 2000).

Inter-rater reliability study: Because the judges could not discriminate reliably enough between the subdimensions Drudgery and Apathy, these two subdimensions were condensed into one which we called Apathy, thus reducing the number of subdimensions to 35.

The mean ICC was 0.82 within a range from 0.54 to 0.89. Using as a standard cut-off score a correlation level of .70, according to the recommendations of Lambert & Hill (1994) for scales that do not have a low level of inference, only 4 of the 35 subdimensions had reliabilities below .70. All of the 35 subdimensions reached Cohen¥s cut-off point of .50 (Cohen, 1988).These remarkable results for inter-rater reliability allowed the continuation of the psychometric investigation and we conducted the extended validity replication study with a new sample, different interviewers and different raters as well as some additional instruments.

Validity study II: Figure 1 shows the mean profile of our group of depressed patients on the 35 SPC subdimensions. This profile proves construct validity of the SPC because it operates in the expected way for depressed patients. The highest mean is for the subdimension Self-Depreciation, followed by Overinvolvement in Relationship, Internalisation, Surrender of Self, and Pessimism (all means are above 1.5).

Insert figure 1 here

Discriminant validity study: The data clearly show no significant correlations between either the
SCL-90-R scales, the BDI, the Impairment-Severity Scores or the two GAF scales and the SPC subdimensions plus the SPC total score (see table 3).

Table 3

Significant correlations (between the 35 SPC subdimensions and the construct-distant / construct-near instruments; $p < 0.001$)

SCL (10 scales)
BDI (1 scale)
BSS (4 scales)
GAF (2 scales)
construct-distant instruments

(r†: .01 - .38)
(r†: .02 - .41)
(r†: .01 - .37)
(r†: .01 - .32)
FPI (12 scales)
IIP (9 scales)
FKBS (5 scales)
OPD (1 scale)
construct-near instruments
12 (r: .46 - .62)
16 (r: .47 - .59)
1 (r = .52)
2 (r = .48 / .49)

Convergent validity study:
As is shown in table 3 there are 12 significant correlations between the 12 FPI scales and the SPC (rr: .46 - .62, $p < 0.001$), and 16 correlations between the IIP scales and the SPC (rr: .47 - .59, $p < 0.001$). Out of the 5 FKBS scales there was only one significant correlation with the scale Turning Against Self (rr = .52, $p < 0.001$). All other defence mechanisms did not correlate. The OPD rating correlated significantly with the SPC total score and the subdimension Coherence (rr = .48 / .49, $p < 0.001$)

CONCLUSIONS

We consider the results of the inter-rater reliability study to be highly satisfactory especially when taking into account that homogenous samples tend to show considerably lower inter-rater reliabilities and that reliability is normally higher when instruments are applied by their developers (Zimmermann, 1994). Seen from an economic point of view, our reliability study suggests that raters do not need to have psychoanalytic training and are doing sufficiently well with a medium range rater training.
The results of the discriminant validity studies clearly reveal that there is no correlation between the SPC and the construct-distant measures, thus confirming our first assumption that the SPC measures beyond symptomatology.
The convergent validity studies with their considerable correlations between the SPC and the construct-near instruments offer convincing evidence that the SPC measures something similar, but not identical to interpersonal functioning and personality, thus confirming our second assumption.

We do not want to go into detail by interpreting single correlations, but nevertheless want to state that all inter-scale relationships were conceptually consistent and meaningful. We would like to point out that the significant correlation between the SPC and the scale Turning Against Self of the FKBS can be interpreted as the Ñpathognomonic introjectionì (Fenichel, 1945) of the depressive patient. This finding can be understood as another test of construct validity. The absence of any other correlations between the SPC and the remaining four scales of the FKBS suggests that the SPC does not measure just defence mechanisms which is in accordance with our working model of ìstructural changeî (Kernberg, 1991, Wallerstein, 1991).

We conclude that from our psychometric studies that the SPC appears to be a reliable and valid instrument. The findings of validity study I were replicated in validity study II, although the interviewers and raters differed considerably between study I and study II in clinical experience, therapeutic education and rater-training.

Although there is still much work to be done to establish other psychometric qualities, above all sensitivity to change, there is already substantial evidence that researchers have available an instrument that measures psychic structure, which is a first step towards the measurement of structural change, the mode-specific effect of psychoanalysis and psychoanalytic psychotherapy.

REFERENCES

Alexander L. B., & Luborsky L. (1986). The Penn Helping Alliance Scales. In: L. S. Greenberg, W. Pinsof (Eds.), The psychotherapeutic process: a research handbook. New York: Guilford Press, 325-366.

American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders. Fourth Edition, Washington DC: APA, 1994.

Bachrach, H. M., Galatzer-Levy, R., Skolnikoff, A., & Waldron, S. (1991). On the efficacy of psychoanalysis. Journal of the American Psychoanalytic Association, 39: 871 ñ 916.

Beck, A. T., Ward, C. H., Mendelsohn, M., Mock, J., & Erbaugh, J. (1961). An inventory for measuring depression. Archives of General Psychiatry, 4: 561-571.

Benjamin, L.S. (1974). Structural Analysis of Social Behaviour. Psychological Review, 81: 392-425.

Cohen, J. (1988). Statistical Power Analysis for Behavioral Sciences. Hillsdale, New York: Lawrence Erlbaum Assoc.

Dahl, H. (1988). Introduction. In H. Dahl, H. Kaechele, & H. Thomae (Eds.), Psychoanalytic Process Research Strategies. Berlin: Springer.

Derogatis, L. R., Lipman, R.S., & Covi, L. (1973). SCL 90. An outpatient psychiatric rating scale. Psychopharmacology Bulletin 9: 13-28.

DeWitt, K. N., Hartley, D., Rosenberg, S. E., Zilberg, N. J., & Wallerstein, R. S. (1991). Scales of psychological capacities: Development of an assessment approach. Psychoanalysis and Contemporary Thought, 14: 334 - 343.

DeWitt, K. N., Milbrath, C., & Wallerstein, R. S. (1999). Scales of Psychological Capacities: Support for a measure of structural change. Psychoanalysis and Contemporary Thought, 22: 453-480.

Elkin, I., Shea, M.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: General effectiveness of treatments. Archives of General Psychiatry, 46: 971-982.

Fahrenberg, J., Selg, H., & Hampel, P. (1989). Das Freiburger Persˆnlichkeits-Inventar (FPI-R). Gˆttingen: Hogrefe.

Fenichel, O. (1945). The Psychoanalytic Theory of Neurosis. New York: Norton.

Grawe, K., & Braun, U. (1994). Qualit‰tskontrolle in der Psychotherapiepraxis. Zeitschrift f¸r Klinische Psychologie, 23: 242-267.

Hamilton, M. (1960). A rating scale for depression. Journal of Neurology, Neurosurgery and Psychiatry, 23: 56-62.

Hentschel, U. (1998). Fragebogen zu Konfliktbew‰ltigungsstrategien. GˆttingenHogrefe.

Hiller, W., Zaudig, M., & Mombour, W. (1995). Internationale Diagnosen-Checkliste f¸r ICD 10 und DSM IV (IDCL). GˆttingenHogrefe.

Horowitz, L. M., Rosenberg, S. E., Bauer, B. A., & Ureno, G., & Villasenor, V. S. (1988). Inventory of Interpersonal Problems: Psychometric properties and clinical applications. Journal of Clinical and Consulting Psychology, 56: 885-892.

Huber, D., Henrich, G., & Herschbach, P. (1988). Measuring the quality of life: A comparison between physically and mentally chronically ill patients and healthy persons. Pharmacopsychiatry 21: 453-455.

Huber, D., & Klug, G. (1997). How to measure structural change. Paper presented at the 28th annual meeting of the Society for Psychotherapy Research, Geilo, 25.-29.6.1997.

Huber, D., Klug, G., & von Rad, M. (1997). M¸nchner Psychotherapie-Studie (MPS). In: M. Leuzinger-Bohleber & U. Stuhr (Eds.), Psychoanalysen im R¸ckblick. Gieflen: Psychosozial.

Huber, D., & Klug, G. (1999). The Munich Psychotherapy Study (MPS) ñ Comparing the effects of psychoanalysis and psychotherapy. In P. Fonagy, H. K‰chele, R. Krause, E. Jones, & E. Perron, (Eds.), An Open Door Review of Outcome Studies in Psychoanalyis. Report prepared by the Research Committee of the IPA at the request of the President. International Psychoanalytic Association, London: http://www.ipa.org.uk.

Huber, D., Klug, G., & von Rad, M. (2000). Die M¸nchner Psychotherapiestudie. Ein Vergleich zwischen Psychoanalysen und psychodynamischen Psychotherapien unter besonderer Ber¸cksichtigung therapiespezifischer Ergebnisse. In U. Stuhr, M. Leuzinger-Bohleber, & M. Beutel (Eds.), Psychoanalytische Langzeittherapien. Stuttgart: Kohlhammer.

Huber, D., Henrich, G., & von Rad, M. (2000). ‹ber den Nutzen von Beratungsgespr‰chen an einer psychosomatisch-psychotherapeutischen Ambulanz. Psychotherapie Psychosomatik Medizinische Psychologie, 50: 147-156.

Jones, E. E., Cumming, J. D., & Horowitz, M. J. (1988). Another look at the nonspecific hypothesis of therapeutic effectiveness. Journal of Consulting and Clinical Psychology, 56: 48-55.

Kernberg, O. F. (1991). Psychic structure and structural change: An ego psychology-object relations theory viewpoint. In T. Shapiro (Ed.), The Concept of Structure. Madison, Connecticut: International Universities Press, Inc.

Kiresuk, T. J., & Sherman,R. E. (1968). GAS: a general method for evaluating comprehensive community mental health programs. Community Mental Health Journal, 4: 443-453.

Lambert, M. J., & Hill, C. E. (1994). Assessing Psychotherapy Outcomes and Processes. In A. E. Bergin & S. L. Garfield (Eds.), Handbook of Psychotherapy and Behaviour Change. New York: J. Wiley & Sons.

Luborsky, L., Diguer, L., Luborsy, E., Singer, B., Dickter, D., & Schmidt, K. A. (1993). The Efficacy of Dynamic Psychotherapies: Is it true that Ñ Everyone has won and all must have prizes?ì In N. E. Miller, L. Luborsky, J. P. Barber, & J. P. Docherty (Eds.), Psychodynamic Treatment Research. New York: Basic Books.

Luborsky, L., Diguer, 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 Ñwild cardì in comparisons of treatment efficacy. Clinical Psychology: Science and Practice, 6: 96-132.

Mercer, R. C., & Loesch, L. C. (1979). Audiotape ratings: Comments and guidelines. Psychotherapy: Theory, Research and Practice, 16: 79 - 85.

Moore, B. E., & Fine, B. D. (1990). Psychoanalytic Terms & Concepts. New Haven: Yale University Press.

OPD-Task Force (Eds.) (2001) Operationalized Psychodynamic Diagnostics. Foundations and Manual. Seattle, Toronto, Gˆttingen, Hogrefe & Huber.

Pfeffer, A. Z. (1959). A procedure for evaluating the results of psychoanalysis: a preliminary report. Journal of the American Psychoanalytic Association, 7: 418 - 444.

Piaget, J. (1970). Structuralism. New York: Basic Books.

Pulver, S. E. (1991). Psychic structure, function, process, and content: toward a definition. In T. Shapiro (Ed.), The Concept of Structure. Madison, Connecticut: International Universities Press, Inc.

Rad, M. von, Senf,W., & Br‰utigam, W. (1998) Psychotherapie und Psychoanalyse in der Krankenversorgung. Ergebnisse des Heidelberger Katamnese-Projektes. Psychotherapie Psychosomatik Medizinische Psychologie, 48: 88 - 100.

Roth, A., & Fonagy, P. (1996). What works for whom? New York: Guilford Press.

Schepank, H. (1995). Der Beeintr‰chtigungs-Schwere-Score (BSS). Gˆttingen: Beltz.

Schulte, D. (1995). How treatment success could be assessed. Psychotherapy Research, 5: 281-296.

Shrout, P. E., & Fleiss, J. L. (1979). Intraclass correlations: Uses in assessing rater-reliability. Psychological Bulletin, 86: 420 - 428.

Sommer, G., & Fydrich, T. (1991). Entwicklung und ‹berpr¸fung eines Fragebogens zur Sozialen Unterst¸tzung (F-SOZU). Diagnostica, 37: 160-178.

Strupp, H. H., Schacht, T. E., & Henry, W. P. (1988). Problem ñ Treatment ñ Outcome Congruence: A principle whose time has come. In H. Dahl, H. Kaechele, & H. Thomae (Eds.), Psychoanalytic Process Research Strategies. Berlin: Springer.

Strupp, H. H., Horowitz, L. M., & Lambert, M. J. (1997). Measuring Patient Changes. Washington DC: American Psychological Association.

Vaughan, S. C., Marshall, R. D., MacKinnon, R. A., Vaughan, R., Mellman, L., & Roose, S. P. (2000). Can we do psychoanalytic outcome research? A feasibility study. International Journal of Psychoanalysis, 81: 513-527.

Wallerstein, R. S. (1986). Forty-two Lives in Treatment. New York: Guilford Press.

.

Wallerstein, R. S. (1991). Assessment of Structural Change in Psychoanalytic Therapy and Research. In: Shapiro, T. (Ed.), The Concept of Structure in Psychoanalysis. Madison: International Universities Press.

Wallerstein, R. S. (O. J.). The Scales of Psychological Capacities. Version 1. unpublished manuscript.

Zielke, M., & Kopf-Mehnert, C. (1978). Ver‰nderungsfragebogen des Erlebens und Verhaltens, VEV. Weinheim: Beltz.

Zimmermann, M. (1994). Diagnosing Personality Disorders. Archives of General Psychiatry, 51: 225-245.

Acknowledgements:

The rater-trainig of the authors (DH and GK) with the PRP-II group in San Francisco at the California Pacific Medical Center has been supported by the German Research Foundation DFG (444 USA 111/3/98).

The inter-rater reliability and the validity study with the SPC are funded by a grant from the IPA.

Close Family or Mere Neighbours? – Some Empirical Data on the Differences between Psychoanalysis and Psychotherapy[2]

Johan Blomberg & Rolf Sandell

Department of Behavioural Sciences, Linköping University

Correspondence concerning this article should be sent to : Johan Blomberg, Samaritgränd 4, 118 53 Stockholm, SWEDEN. e-mail: jo.blomberg@telia.com.

[2] This article is an extended version of a paper presented at the 17th World Congress of Psychotherapy, Warsaw, Poland, August 23-28, 1998.

A General Background

Ever since Freud's times there has been discussion about the differences and similarities between psychoanalysis and psychoanalytically oriented therapies. The questions discussed are theoretical, practical, and political in nature: Are the differences in indications, technique and processes mainly a matter of *degree* or of *quality*, the latter being a stricter distinction? Is it possible to develop a genuine psychoanalytic process in once- or twice-a-week treatment? Are psychoanalysts automatically qualified to do and teach psychotherapy? Should psychotherapy be taught as part of the psychoanalytic training ?

The degree of segregation among the treatment modalities in the psychoanalytic part of the psychotherapy spectrum seems to vary between different countries. In Sweden psychoanalysts and psychoanalytically orientated therapists are quite careful – some would say meticulous - about the distinctions.[3] In Germany, on the other hand, the distinction is less clear and the expanded term "intensive psychotherapy" (Dossmann, Kutter, Heinzel & Wurmser, 1997) seems to be used to refer to a grey zone between psychoanalysis proper and low frequency psychotherapy on a psychoanalytic footing. In this paper we will simply use the terms *therapy* and *therapist* as generic terms when we refer to the entire spectrum of psychoanalytic treatments and the terms *psychoanalysis* and *psychotherapy* when we refer to psychoanalysis and psychoanalytically oriented psychotherapy specifically.

In 1954 the American Psychoanalytic Association set up four full-day conference panels on the differences and similarities between psychoanalytically oriented psychotherapy and psychoanalysis proper. In the discussions two groups of participants emerged, those who viewed psychoanalysis and psychoanalytic psychotherapy as distinctively separate modalities (e.g., Bibring, 1954; Gill, 1954; Rangell, 1954) and those who would blur the boundaries or see none at all (e.g. Alexander, 1954; Fromm-Reichmann, 1954). Today, almost 50 years later, it seems as if the issues discussed and the groupings are pretty much the same, Kernberg (1999), for example, being in favour of clear distinctions, and Fosshage (1997) favouring no clear distinction[4].
From an empirical point of view there is indeed no evidence leading to a strict distinction between psychoanalysis and psychotherapy. To this day, the most ambitious project making a relevant comparison seems to be the Psychotherapy Research Project (PRP) of the Menninger Foundation, launched as early as in the 50's. Based on the accumulated findings in the PRP, Wallerstein concluded clearly in favour of those who would blur the

[3] There has been a sometimes heated dispute between the two Swedish societies as to whether psychoanalysis could be conducted three times a week or whether this, by definition, should be called psychotherapy.
[4] It is interesting to note that, with few exceptions, it is psychoanalysts who discuss these issues.

boundaries: "The therapeutic modalities of psychoanalysis, expressive psychotherapy, and supportive psychotherapy hardly exist in ideal or pure form in the real World of actual practice. /…/ (treatments) are intermingled blends of expressive-interpretative and supportive-stabilising elements…. and /…/ the overall outcomes achieved by more analytic and more supportive treatments converge more than our usual expectations for those differing modalities would portend; and the kinds of changes achieved in treatment from the two end of this spectrum are less different in nature and in permanence than is usually expected." (Wallerstein, 1989, p. 205). Thus, contrary to what was expected, there were no differences in outcomes after psychotherapy and psychoanalysis; the mean effects of either treatment were quite modest; supportive techniques were as powerful as more interpretative ones; and psychoanalysts used supportive techniques to a larger extent than what was usually assumed.

We believe that the findings of the Menninger study have been vitalising to the discussion on the psychotherapy versus psychoanalysis issue by putting some empirical "facts" in focus. There is a need for more such empirical data. Unfortunately, there has been a strange disinterest among psychoanalysts, psychotherapists and, indeed researchers, in collecting systematic data on long-term psychoanalytic psychotherapy and psychoanalysis. Further, the quality of the few systematic outcome studies that have indeed been undertaken (as reviewed by Bachrach, Galatzer-Levy, Skolnikoff & Waldron, 1991; Doidge, 1997; Kantrowitz, 1997) has generally been poor (Fisher & Greenberg, 1996). At the same time, research on process and outcome in short-term psychoanalytic treatment abound. Apart from making psychoanalysis and long-term psychotherapy vulnerable to attacks from adherents to so called empirically validated (usually short-term) treatments, the lack of empirical data has tended to transform important theoretical discussions into introvert academic hair-splitting, with little or no impact on training or practice. Humbly, we hope that the facts presented here will contribute to further the discussion.
In 1993 we launched the Stockholm Outcome of Psychotherapy and Psychoanalysis Project (henceforth, STOPP). In this paper we offer a condensed account of some of the results and, for the first time, present a questionnaire, The Therapeutic Attitudes Scale (TASC) which consists of three different scales, mapping what we like to think of as the *therapeutic milieu* which therapists provide for their patients. We have tried to avoid heavy technical and methodological descriptions and instead focus on results that we believe may challenge some of the conclusions drawn from the Menninger project.

Our approach will be as follows. We begin with the background and the basic design of STOPP. Then, we will present data on patient characteristics and the general outcome. We will then introduce the Therapeutic Attitudes Scale (TASC) and some empirical differences between psychotherapists and psychoanalysts with respect to what they

believe and what they do, or at least claim to do, in their offices. Finally, we present some data and a discussion on how the beliefs and attitudes of the treatment providers interact with treatment modality, producing differential effects in psychotherapy and in psychoanalysis.

Where We Started – The Study Background

In 1989, the Swedish national health insurance authorities decided to specially fund projects to alleviate illnesses that were particularly burdensome to the national health insurance and health care systems. Among the different projects funded, one was for psychoanalysis and long-term psychotherapy with therapists in private practice. In Sweden, therapy is covered by the national health insurance system only as long as a medical doctor provides it. However, a majority of the licensed therapists are not physicians. Since many people do not have the possibility, or are not prepared, to pay for long-term intensive treatment, we have, in effect, a situation with long waiting lists for therapy with medical doctors and an under-utilised capacity of therapy with non-physicians. The STOPP project was launched with grants from the national health insurance authorities. The main purpose was to study which people, under what circumstances, seek psychoanalysis or psychoanalytic psychotherapy and what benefits they were able to derive from their treatments.

A program supervisor decided on the subsidisation of a treatment. To have subsidy the patient should first have contracted a licensed therapist and then have a written referral from another therapist, with a description of his or her need, and suitability for the suggested treatment. The subsidy covered all costs up to three years with no extensions allowed. (However, patients were free to continue treatment beyond the three years financing it in other ways.) One effect of this quite bureaucratic procedure, clearly not intended, was that most patients' treatments had already started at the time of the referral. Also, we strongly suspect that in most cases the referring therapist made only a passive judgement of whether the person could benefit from the suggested treatment, rather than an independent assessment of what he or she believed was the treatment of choice for this particular patient. Hence, we believe that to a large extent we are dealing with "ordinary", usually highly motivated, patients who have actively chosen both their therapist and their mode of treatment.

What We Did and Why – The Basic Structure of the Study

The project involved almost 1 200 patients in all. Our sample consisted initially of 756 persons: the 202 who had received subsidy between the years 1991-1993 and the first 554 on the waiting list for subsidy at the time of the first of three waves of follow-up, in 1994.

First, we read all the referrals in order to map what kind of patients we were dealing with. As it turned out many patients were already in treatment at the time of the referral, assessment of pretreatment status had to be based first and foremost on the referrals. To complicate further, diagnoses had to be made retrospective for those who had already begun treatment. - What was the case when the patients started? Fortunately, most referrals described patients history in such a way as to make this possible. However, exact diagnoses could not be done with such database, so each patient was only grossly diagnosed by a research assistant (psychologist) as having or not having a DSM-IV axis I or II diagnosis (American Psychiatric Association, 1994). We also assessed each person according to the Global Assessment of Functioning scale (GAF; DSM-IV, axis V). Besides current state, based on the descriptions we also made a rating of the lowest level of functioning after age 18, which is an invention for this study. The reliability of our diagnostic efforts where checked with three judges making three independent diagnoses on a subsample of 20 patients. Despite the difficulties, the interrater reliabilities were found to be acceptable to very good: *ICC* = .69 for the presence of an axis I diagnosis and .51 for the presence of an axis II diagnosis, and .69 and .88 for current and lowest GAF, respectively.

Two postal questionnaires were the basic motors in our design. (a) The *Well-being Questionnaire* (*WbQ*), which was sent to all the patients in May 1994, 1995 and 1996, and (b) the *Therapeutic Identity* (*ThId*), which was sent to all treatment providers, i.e. therapists and analysts, in the spring of 1996.

The WbQ is a 24-page booklet, specifically designed for this project. It contained a series of questions and items focusing on demographic, familial, and socio-economic conditions; data on frequency and duration of ongoing or terminated treatment(s); previous treatments; sickness and health care utilisation. It also included a number of well-known self-rating scales: the Symptom Check List-90 (SCL-90; Derogatis, Lipman, Rickels, Uhlenhuth & Covi, 1974); the Sense of Coherence Scale (Antonovsky, 1987), and the Social Adjustment Scale (Weissman & Bothwell, 1976), all of which have high or very high reliabilities (between .60 and .90 in different studies). In this paper we will concentrate on the SCL-90 only[5]. The SCL-90 consists of 90 items for each of which the patient is asked to check on a five-point scales ranging from 0 (*not at all*) to 4 (*very much*), to what extent he or she has been troubled by various signs of somatic and psychic distress during the last seven days (for references on the reliability and validity of SCL-90 see Bridges & Goldberg, 1989).

[5] These three scales were highly correlated which means that they either measure the same thing, or that they are dependent on eachother.

After three reminders about 60% of the sample had remained in the panel through all three waves. In this paper we will present data from 331 persons who were in, or had been in, long-term psychotherapy, and 74 persons whose main treatment was, or had been, psychoanalysis.[6] An analysis of the attrition revealed only small and typical differences between responders and non-responders. Those who did not want to partake in the study at all had lower current GAF before treatment, and those who dropped out of the study had lower levels of education and lower level of functioning according to current GAF (*p*'s < .05; two-tailed t-tests and chi-squares). However, there were no differences in response rates among patients in psychotherapy versus psychoanalysis.

The ThId was distributed in 1996 to all 316 treatment providers (therapists and analysts) who had at least one patient (the range was 1-11 patients) in the project. For norming and standardising purposes the ThId was also sent to a random sample of 325 Swedish therapists. The ThId has about 150 questions and/or items, divided in different sections. Apart from the TASC, which will be described in detail later on, there were three sections dealing with (a) basic education and professional training; (b) professional experience; and (c) personal therapy or training analysis.

It took four reminders to get 227, or 69%, of the national sample and 209, or 66%, of the 316 psychotherapists and psychoanalysts in the STOPP sample to complete the questionnaire. Chi-square analyses showed that attrition was not systematic (all *p*'s < .05), except for the fact that therapists over 65 years of age in the national sample tended to abstain from responding on account of their retirement.

Who Seek Psychotherapy and Who Seek Psychoanalysis?

Psychoanalytically orientated psychotherapists and psychoanalysts are sometimes accused of treating only "the worried well". There is a widespread notion that the typical patient in psychoanalytically oriented treatments is a "YAVIS" (Young, Attractive, Verbal, Intelligent, and Social), that is, well-off Woody-Allen-like characters whose main problem is an addiction to therapy. But when one review studies on consumers of psychotherapy and psychoanalysis, one will find that the typical patient bears only slight resemblance to this picture. In fact, she is an urban, middle-aged, well-educated, professional female, typically in a health-care, educational, or artistic profession (Carlsson, 1991; Carlsson, 1993; Garfield, 1994; Olfson & Pincus, 1994 b; Vessey & Howard, 1994; Weber, Solomon & Bachrach, 1985). At the same time, however, she is typically in a fairly bad shape in terms of her general medical condition, health-care

[6] Twenty persons were excluded because of incomplete data and 12, highly unrepresentative persons who never started their treatments were excluded and 13 patients in various kinds of low-dose therapies, viz. brief, low frequency, group, or family therapies were excluded.

consumption, and capacity to work. Indeed, this seems to be the case both in Sweden (Carlsson, 1991; Carlsson, 1993; Schubert & Blomberg, 1994), in the United States (Olfson & Pincus, 1994 a; Olfson & Pincus, 1994 b; Doidge, Simon, Gillies & Ruskin, 1994), and in Germany (Dossman, Kutter, Heinzel & Wurmser, 1997).

In analysing the referrals we found that, in line with earlier findings, the patients in our sample appeared as a highly qualified group, educationally and vocationally. They were clearly not a representative sample of the general population, apparently rather belonging to the cream of the crop. However, based on the coarse pre-treatment diagnoses and some complementary pre-treatment data from the WbQ, we found that the patients were quite vulnerable and often highly distressed, with long self-reported histories of suffering and histories of psychiatric (mainly out-patient) care and use of psychoactive drugs. The typical patient had had psychiatric problems for more than five years, and the most frequent symptom was feelings of self-blame and stress for not having things done properly or on time (items 26 and 86 on the SCL-90).

Table 1 shows the differences between patients in psychotherapy and psychoanalysis, respectively. Although they are similar in some respects, there were also some significant differences. As can be seen, there were relatively more men among the analysands. Analysands were also a few years older, had higher levels of education (which were already high among the psychotherapy patients), were more often married or divorced and had more frequently children. Psychiatrically, there were no, or only small, diagnostic differences or differences with respect to syndromes or level of disturbance according to the GAF scales. However, before the present treatment, psychotherapy patients tended to have utilised institutionalised psychiatry (inpatient or outpatient clinics, emergency rooms etc.) to a larger extent, whereas analysands had rather turned to psychotherapy for help.

Table 1. Sociodemographic and Diagnostic Breakdown of Analysands and Psychotherapy Patients.

		Psychotherapy Patients (n = 331)	*Analysands* (n = 74)	
Sociodemographic Characteristics				
Men	%	20	37	**
Age	*M*	36	40	**
Married and/or divorced	%	38	60	**

Cohabiting with a partner	%	45	47	
Has children	%	49	66	*
Have some college education	%	76	94	**
DSM-IV Categories				
Psychiatric syndrome (axis I)	%	58	54	
Personality disorders (axis II)	%	12	11	
No psychiatric diagnose (V-codes)	%	33	36	
Level of Functioning				
GAF, current	*M*	60	61	*
GAF, lowest after 18 years of age	*M*	52	54	
Previous Psychiatric Treatments				
Any psychiatric treatment at all	%	79	91	*
Psychotherapy	%	63	75	*
Psychoactive drugs	%	55	56	
Outpatient psychiatric care	%	56	45	
Psychiatric emergency room	%	38	26	
Hospitalized	%	21	10	*

***Note*. * *p* < .05; ** *p* < .01, by *Chi-square* and two-tailed *t*-tests for differences between the two modalities.**

We conclude that, despite great similarities, there are some interesting differences between persons how seek psychotherapy and those who seek psychoanalysis. It seems that socio-economic and socio-cultural factors are the most important ones to distinguish between the two groups: The higher the social status of the patients, the more likely he or she is in psychoanalysis. One may only speculate on why this is the case. In general, highly educated people, especially in the social, educational, and health-care sectors or in the humanities, are familiar with, and take an active interest in, psychoanalytic thinking. They are, as Kadushin (1969) puts it, the "friends and supporters" of (psychoanalytic) therapy. The difference also mirrors the status or prestige differential between psychotherapists and

psychoanalysts. Is it the patients or the therapists who generates it? Who chooses whom? Several studies suggest that the relationship between the social status of the psychotherapist and his or her patients indeed has to do with the fact that high status therapists tend to choose high status patients (Garfield, 1986; Kadushin, 1969; Lubin, Hornstra, Lewis & Bechtel, 1973; Weber, Solomon & Bachrach, 1985). No doubt, the social selectivity is a political problem insofar as the treatment is paid for by public money. In Sweden the attacks on psychoanalysts and psychotherapists for their alleged lack of social concerns are frequent and, sometimes, harsh. Biological psychiatry and the cognitive-behavioural treatments maintain an attitude of being more attuned to "ordinary" or "lower- class" patients -- by implication those who "really suffer". As psychoanalysts and long-term psychotherapists lack empirical data, they tend to withdraw into a defensive position; sometimes claiming that the severe psychiatric patients are not really suitable for psychoanalytically orientated treatments. For one thing, this does indeed not seem to be the case, as far as our data show, for another, this offers an advantage to other kinds of treatments in the battle for public subsidy. The "worried-well" accusation does really not seem to be fair. Our interpretation is that it is the "burnt-out" and distressed professionals who seek therapy rather than bored housewives and artists looking for pastime. A majority of them have tried different treatments, including drugs, before entering long-term treatment. Also, many had tried other types of, usually less intensive, psychotherapy before. A related question of great interest is whether retakes or new treatment are made less probable by long-term treatments? In fact, some findings of ours suggest that this may be the case (Blomberg, et. al, 1997, June).

Did Patients get any Better? – The Treatment Effects

In order to answer this question we had to partition our two treatment groups in such a way as to allow some form of comparison between subgroups before, during and after treatment. Remember that many of the patients had already started treatment when we started our project, because they had not waited for subsidy but started whenever they had found any way to finance it. Hence, we had no control over when patients in fact started or terminated treatment. Indeed, the timing of our follow-up questionnaire (the WbQ) was totally independent of where any one particular patient was in his or her treatment process. What we had, then, was a pool of patients in different phases of their treatment processes. Consequently, our first follow-up questionnaire "hit" the patient randomly with respect to where he or she was in the treatment process, before, during, or after treatment. Figure 1 is an attempt to illustrate this pictorially.

Figure 1. Examples of Patients at Different Stages of Treatment, in Relation to Waves of Administration of the Well-being Questionnaire in 1994 (WbQ1), 1995 (WbQ2) and 1996 (WbQ3).

Consider three persons, A, B, and C, with treatments extended in time as indicated by the thick arrows. At the time of our first administration of the WbQ, patient A had not yet begun her treatment, whereas patient B was in treatment and patient C had already finished hers. At the second administration (WbQ 2), patient A had recently begun treatment, patient B was a little bit further in hers, whereas patient C was one year further post-treatment. At the time of the third administration (WbQ 3), patient A was still in treatment, at a later stage in her treatment, whereas patient B had now terminated treatment. Patient C was yet another year further in her post-treatment process.

The story of *Alice's Adventures in Wonderland* may be used as an analogy for how we designed our data-analyses[7]. At one point in the story, Alice had wept violently, creating a pool of tears crowded with birds and animals that had fallen into it. The Dodo then suggested a "Caucus-race" to get them dry: "First it marked out a race-course, in a sort of circle (the exact shape doesn't matter, it said), and then all the party were placed along the course, here and there. There was no 'one, two, three, and away,' but they began running when they liked, and left off when they liked, so that it was not easy to know when the race was over. However, when they had been running half an hour or so, and were quite dry again, the Dodo suddenly called out 'The race is over!' and they all crowded round it, panting, and asking 'But who has won?' (Carroll, 1865, p. 32-33). Using this analogy, in our "race", instead of the "race-is-over" call-out, we had three intermediate checkpoints (where we administered the WbQ), with one-year intervals. At each of these checkpoints we measured the patients' well being (just as the Dodo might have measured the animals' pulse, if he'd been interested enough). At any of these checkpoints any one patient then could be before, during, or after treatment. After collecting all data we grouped each of the 1 250 checkpoint scores or observations according to where the patient was in his or her treatment at the time, before, during or after treatment and, more specifically, in various subdivisions of each. In effect, we were able to

[7] This story is often used to illustrate another phenomenon in psychotherapy research, namely the so-called Dodo bird verdict, according to which "all (therapies) have won and all must have prizes".

position each checkpoint observation in one of the following seven phases on a relative time scale[8].

Table 2. The Number of Observations at Each of the Seven Treatments Phases on a relative Time Scale (N = 1250).

Phase	No. of observations
1. pre-treatment [9]	**35**
2. early treatment	**186**
3. mid treatment	**207**
4. late treatment	**227**
5. early posttreatment	**232**
6. mid posttreatment	**207**
7. late posttreatment	**156**

In analysing the SCL-90 scores we simply calculated and plotted the mean scores across all observations in each position along the seven-point time scale. Figure 2 shows a decay curve for each treatment group[10].

[8] Technically, we had a quasi-experimental design that is partly cross-sectional (across stages of treatment; across treatment modalities) and partly longitudinal (across successive stages of treatment).

[9] In principle, we had indeed *three* phases before treatment, one with a number of observations on patients who never started any treatment at all and one with observations only on a small number of patients who were in other kinds of treatments than psychoanalysis or psychotherapy. These observations and, accordingly, these phases, were therefore discarded.

[10] In technical terms we have created a partly within- and partly between subjects design which we consider this equivalent to randomisation since our follow-ups was random with respect to patients treatment processes (Chambless & Hollon, 1998).

Figure 2. SCL-90: Decay Curves: Mean SCL-90 scores for Patients in Psychotherapy and Psychoanalysis Across Different Phases of the Treatment Process.

As can be seen, the analysands and psychotherapy patients started off at almost identical levels of symptom distress before treatment and both groups then got better at about equal pace. However, after treatment termination the analysands got progressively better, whereas the mean outcome flattened out asymptotically after psychotherapy. Analysing the linear trends the intercepts and *b coefficients* were both significantly different from 0 in both groups, t (6) = -8.07, p < .001. Also, the difference between the slopes in the two groups was significant, t (12) = 4.08, p < .010. In terms of so-called effect size (*d*) the difference between the first and last values in the series was 0.59 for psychotherapy and 1.55 for psychoanalysis. According to research conventions *d*'s between 0.50 and 0.75 are moderate and over 1.00 are large. Hence, the effect size for analysis is *very* large when compared to the effect sizes in the oft-cited meta-analyses of, for example, Smith, Glass and Miller (1980) and Shapiro and Shapiro (1982).

Now, how do we know that these differences do not stem from differences in patient characteristics? In fact and as already noted, the patients were indeed not very different in absolute terms, in the first place. However, to make sure, using regression-analyses, we partialled out diagnostic and demographic factors statistically and found that this did not change the differential trend.

More important, one might ask whether these differences were caused mainly by "extrinsic" factors such as treatment dose (frequency x duration), by the providers' formal competence, or mainly by "intrinsic" factors such as the therapeutic techniques used and the therapeutic processes that developed? There were indeed great differences with respect to the number of sessions in the two groups. The psychoanalyses were three to five times per week (*M* = 3.6; *SD* = 0.7) with a member of one of the psychoanalytic societies in Sweden (one within the IPA and one, at the time, within the IFPS), and the psychotherapies were once- or twice-a-week (*M* = 1.5, *SD* = 0.5) with a psychotherapist licensed by the National Board of Health and Welfare. At an average then, the psychoanalyses had a total of 642 sessions (*SD* = 324) over about four and a half years (*M* = 54 months; *SD* = 23) and the psychotherapies 233 sessions (*SD* = 151) over nearly four years (*M* = 46 months; *SD* = 24).

The obvious question now is whether these extrinsic time factors, frequency, duration and dose (the total number of sessions) account for the differences in effect we observe? We have tried to sort this out in great detail using a sophisticated statistical procedure known as structural equation modeling (Jöreskog & Sörbom, 1986; Kline, 1998). It would bear too far to give an account here. However, we found the time factors *alone* could not explain the differences: neither frequency, nor duration had any effects in and by themselves or separately, whereas their interaction could explain *some* of the outcome variance (Sandell, Blomberg & Lazar, 2000). Since patient variables could not explain the variance, we believe that other variables must be at work accounting for at least some of the differences. Next, we will turn to one of these.

Different Folks provide Different Strokes! Differences and Similarities among Treatment Providers

One important lesson from the Menninger project is that, in comparing psychotherapy and psychoanalysis, one cannot rely on any notion of a "standard" psychoanalytic or psychotherapeutic technique, since psychoanalysts and psychotherapists vary considerably in how they actually practice, despite similar orientations and training. Clearly, one main problem with our study was the lack of *direct* observations — through recordings or notes—of what had really been going on in treatment.

Of course, we might have tried to collect retrospective data from patients and therapists, but such retrospective data are indeed problematic, both from a methodological and an ethical perspective. Our way of handling this was to develop a set of measures on therapist variables that we thought would be *related to* what might have been going on in the treatments. Of course, one can not

assume any one-to-one correspondence between words and deeds, and a therapist's technique will probably vary somewhat with the patient. However, it was—and is—our assumption that the therapist's beliefs and values in therapeutic matters help determine the general approach and technique he or she actually uses. Also, several studies indicate that therapists intentions are in fact as strong, or even stronger, predictors of outcome than the their actual deeds (Hill, 1988; Taylor, Adelman & Kayser-Boyd, 1986). If we think of the consulting room as a therapeutic "black box", one could say that we have not been able hear exact what they are saying in it, but we can hear the tone of their voices.

The inventory we created is called the Therapeutic Attitudes Scale (TASC) and is part of the larger Therapeutic Identity questionnaire (ThId). The TASC consists of three sets of subscales:

- Curative factors
- Therapeutic style
- Basic assumptions

The *Curative factor* scales were based on 33 items, initially, to rate the therapist's beliefs in the curative value of each of a number of ingredients of therapy (e.g., Helping the patient avoid anxiety-provoking situations). The instruction ran as follows, "What do you think contributes to long-term and stable therapeutic change?" and the rating of each item was made on a five-step scale, from 0 (*not at all*) to 4 (*a lot*).

The *Therapeutic style* scales were based on 31 original items to describe the therapist's own manner of conducting psychotherapy, in the general case ("What are you like as a therapist?"). Again, the items (e.g., I keep my personal opinions and circumstances completely outside the therapy) were rated on a five-point scales, from 0 (*not at all*) to 4 (*a lot*).
The *Basic assumptions* scales were based on an initial series of 16 items relating to one's more basic assumptions about the nature of psychotherapy and the nature of the human mind, partly inspired by Hjelle and Ziegler (1981). The rating scales were continuous bipolar scales, with anchors at each of the poles offering a completion of the item stem (e.g., Psychotherapy may be described ... as a science/as a form or art). The ratings were later transformed to five-point scales. To encourage wider use of the scales, the scales and their items are presented in the Appendix.

When we analysed the TASC-scales we found several important and significant differences within the sample. Figure 3 shows the relative difference between the treatment providers. As can be seen, patients in psychotherapy, (which were indeed sometimes conducted by people with psychoanalytic training), were treated in a therapeutic milieu that was more similar to what we had found characteristic of behavioural and cognitive therapists in the national sample. The psychotherapy providers

did put significantly greater value than the psychoanalysis providers on promoting *adjustment* and showing *kindness* as curative factors (*d*'s > 0.43, *p*'s < .05). Psychotherapy providers also preferred a technique that was less *neutral* and higher on *supportiveness* and s*elf-disclosure* (*d*'s > 0.26; *p*'s < .05) than the psychoanalysis providers. However, there were no significant differences with respect to the basic assumptions, although those providing psychoanalysis were a little bit higher on *pessimism* and *irrationality* (*d*'s > 0.23; n.s.).

*** = *p* < .05.**

Figure 3. Mean z-scores on the TASC for Treatment Providers in Psychoanalysis (broken line) and in Psychotherapy (unbroken line). Curative factors (Adjustment, Kindness and Insight), Therapeutic style (Supportiveness, Self-disclosure, Neutrality), and Basic Assumptions (Art, Irrationality, Pessimism).

Based on the scores on all nine sub-scales (using Cluster analyses) each therapist was then assigned to one of the four standard clusters, based on the national sample, described in the Appendix. The distribution was as follows:

- 42% were assigned to psychoanalytic cluster
- 34% and 24%, respectively, were assigned to the eclectic clusters
- 0% to the cognitive/cognitive-behavioural cluster.
-

Thus, if we are to believe that treatment is carried out in rough correspondence with what the treatment provider believes and values—or at least claims to believe and value, we find that neither psychoanalysis, nor

psychotherapy, are unique or specific designations of any particular technical approach. The division of the therapists in three—or at least two—distinct clusters is all the more interesting, in light of the fact that, in response to a direct question in the ThId, 95% of the therapists in our sample claimed to endorse a "psychoanalytic orientation". That psychoanalytically orientated psychotherapists as well as psychoanalysts hold insight in high esteem and value neutrality as a technical attitude is, of course, not astounding—although the value of neutrality may not be as great, generally, in conducting psychotherapy as in psychoanalysis. Over and above sharing these basic tenets, however, therapists will vary in a number of other respects that we assume will make an essential difference in the treatment office. This is news, we believe. If we characterise the treatments on the basis of the therapist's cluster membership, there was a minority (30%) of eclectically orientated psychoanalyses (and hence 70% classically psychoanalytic). Among the psychotherapies, there was a more even distribution of treatments conducted in an orthodox psychoanalytic vein (43%), or in a more eclectic way (57%), with attitudes more characteristic of cognitive and cognitive-behavioural therapies.

Thus, what we have found quite frequently among so-called psychoanalytically orientated psychotherapists, and not infrequently among psychoanalysts, is a particular constellation of attitudes that we would consider as a more sociable, if not humane, attitude in relating to the patient—and which is partly at odds with the neutral attitude—or with a generally psychoanalytic attitude. In conjunction with the positive valuation of insight and neutrality, these attitudes reflect what we have chosen to call an eclectic type of approach which we would hypothesise is more closely connected to what is usually described as the aims of psychoanalytic psychotherapy than to those of psychoanalysis proper.

The Relationship between Therapeutic Milieu and Outcome

Now, the obvious question is "Which type of therapist did best?" Although this question may seem simplistic, few therapists would doubt that the personal and professional qualities are of great importance in contributing to good or not-so-good treatment results. Also, most training institutes put heavy weight on "suitability" for becoming a therapist. However, there is little or no empirical support for any given characteristic, even though therapist expectations and qualities are among the most frequently studied variables in outcome research. In summarising this research, Beutler, Machado and Neufeldt (1994) draw three conclusions: (1) Individual therapist variables account for more variance in outcome than mode of therapy; (2) some therapists are better than others; and (3) some therapists are detrimental to clients. However, apart from this, the relationships between outcome and therapist variables are highly unclear: There is little or no consensus as to which traits of therapists are predictive of outcome.

And, according to the well-known Dodo bird verdict, theoretical orientation *per se*–-in terms of the usual coarse distinctions we make, psychodynamic, cognitive, or behavioural, are obviously not predictive of outcome.

In order to study the relationship between therapists' values and patient outcome we paired all patients with their respective therapists. We then had data from 330 therapist-patient couples, 55 in psychoanalysis and 275 in long-term psychotherapy.

When the three clusters were compared on the basis of outcomes among their patients, we found that, in general, therapists providing an orthodox psychoanalytic milieu did significantly *worse* than what those providing a more eclectic milieu did. Also, when we compared therapists high (above the median) on single TASC scales with therapists low (below the median), we found significant differences in favour of attitudes on which the members of the psychoanalytic cluster were low. Thus, the strongest associations with good outcome were found with therapists high on kindness and supportiveness, neither of which were favoured by the orthodox attitude; art, favoured by neither the orthodox nor the eclectic attitude; and neutrality, favoured by the orthodox view, all p's < .05. Clearly, these findings came as a surprise, since they were rather at odds with the fact that patients treated in psychoanalysis did better. How could this be? The riddle got its answer when we distinguished between the psychotherapy and the psychoanalysis cases and assigned the cases of each type to one of the two main groups, those who had been offered an eclectic milieu and those treated in an orthodox psychoanalytic milieu. The former group was formed of those cases where treatment was provided by therapists in any of the two eclectic clusters and the latter of cases treated by therapists in the orthodox analytic cluster. Figure 4 shows the results.

Figure 4. SCL-90: Decay Curves for Patients Treated in an Orthodox Psychoanalytic Milieu or in an Eclectic Milieu: Mean SCL-90 Scores for Patients in Psychotherapy (unbroken lines) and Psychoanalysis (broken lines).

Obviously, for patients treated in psychotherapy the two types of therapists did significantly differently. Psychotherapy provided in an orthodox psychoanalytic milieu had significantly worse outcomes than psychotherapy in an eclectic milieu. In fact, the mean change among the former cases did not differ significantly from 0, b = -0.007, t (6) = 0.33, n.s. In the eclectic group, b = -0.101, p = .000. Among the psychoanalysis cases there was indeed a difference

in the same direction but not at all significant, b = -0.140, p = .006 versus b = -0.087, p = .012. Thus, what seems to make a *large* difference in psychotherapy, does not (as much) in psychoanalysis, apparently.

Closing Comments

There currently seems to be a trend towards blurring of the boundaries between psychoanalytic psychotherapy and psychoanalysis proper. The lack of data supporting a clear distinction, the multiplicity of psychoanalytic theories, and the pluralism and relativistic trend within psychoanalysis (Fosshage, 1997), all work in this direction.

Although there are indeed great similarities between the two, our results indicate that there are also some very important differences between psychoanalytic psychotherapy and psychoanalysis: patients, outcome, therapist attitudes and techniques. The most interesting finding, however, is that a strict psychoanalytic attitude does not seem to be really appropriate in psychotherapy. The "as-if psychoanalyses" are clearly not successful. This is, we believe, a very important finding that may justify the theoretical discussions.

However, it seems that the theoretical discussions on differences and similarities very often are based on the (explicit or implicit) assumption that a psychoanalytic process is always a good one. Often, it seems that the proponents of a clear distinction as well as those against such a division end up drawing similar conclusions. For example, Gill (1988), a "separatist", closes a discussion on the conversion of psychotherapy into psychoanalysis, thus: "The question of converting psychotherapy into psychoanalysis should rarely arise in the practice of a psychoanalyst because almost always he should be practising psychoanalysis" (p. 262). Fosshage (1991), from the other position, ends up concluding that "What's critical is not the differentiation between psychoanalytic psychotherapy and psychoanalysis, but the consistent application of expanded psychoanalytic technique within the work that we do as psychoanalysts, /…/. In this sense, psychoanalytic psychotherapy cannot be substituted for (or, I will add, converted into) psychoanalysis, it *is* psychoanalysis." (pp. 70-71). No matter the starting point or line of argument, then, there seems to be a norm where "the more psychoanalytic the better" !

Our results challenge such a view. Gill (1954, 1984) has suggested two definitions of psychoanalysis. The "intrinsic criteria" are, analysis of transference, a neutral analyst, the induction of a regressive transference neurosis and the resolution of that neurosis by interpretation; whereas the extrinsic criteria are, "frequent sessions, the couch, a relatively well integrated (analysable) patient/…/, and a fully trained psychoanalyst." Indeed, there have been many articles dealing with the question of whether a psychoanalytic process (the intrinsic criteria) can take place when

extrinsic criteria are not met. However, we want to approach the question from a slightly different point-of-view.
Assuming a relationship between what we have called "the therapeutic milieu" and Gills' intrinsic criteria, the interaction between therapeutic milieu and treatment modality could be interpreted to mean that an orthodox psychoanalytic attitude is indeed sub-optimal or even dysfunctional in a setting which is not psychoanalytic. Put another way: Both the intrinsic and the extrinsic criteria should be met in order for psychoanalysis to be successful. It may be, that a psychoanalytic process, as fostered by an orthodox neutral, interpretative and transference-focused stance of the analyst, may become destructive, mystifying, or confusing to the patient when the extrinsic criteria are not met. The primary question, then, is not whether a genuine psychoanalytic process *can* take place when extrinsic criteria are not met, but rather whether it is good that it does ?

We suggest that, instead of thinking of psychoanalytic psychotherapy as a diluted, or second-rate, form of psychoanalysis, it is more productive to think of it as a unique way of treating patients. Even if it relies on the psychoanalytic theories about pathogenesis and about psychic life in general, psychoanalytic psychotherapy may find its future in openness to other bases of knowledge *as well* as the psychoanalytic, when therapeutic technique is concerned. This does not necessarily mean that psychoanalytically oriented psychotherapists (or, for that matter, psychoanalysts doing therapy) should adopt whatever is in fashion, but rather that they should develop an openness to the multiplicity of therapeutic schools currently at hand. We agree with Kernberg's (1999) view that an active experimentation with different techniques paired with critical systematic research is the best basis for providing fruitful technical advances in both psychoanalytic psychotherapy and psychoanalysis proper.

This leads to the question of whether psychoanalysts are automatically qualified to do and teach psychotherapy? Our answer would be that there are good reasons for psychoanalytic institutes to teach psychotherapy to their candidates. Many psychoanalysts have nowadays a hard time trying to earn a living on doing psychoanalysis only. Therefore they do psychotherapy, instead. It is important that they realise that they should then not engage in some kind of diluted, "as-if" psychoanalysis. Also, in Sweden, psychoanalytically oriented therapists are usually trained and supervised by psychoanalysts. Based on our findings, the sometimes elevated status of psychoanalysts in Swedish training institutes for psychotherapists could be questioned. Again, we agree with Kernberg (1999): a clear delimitation between psychoanalysis and psychoanalytic psychotherapy could help provide a good basis for a broadening and deepening of the understandings of curative factors and the different pathways to therapeutic change. We conclude that courses in psychotherapy within the analytic institutes could enrich both the analysts' educational experience

and contribute to a mutually productive exchange between psychotherapists and psychoanalysts.

References

Alexander, F. (1954). Psychoanalysis and psychotherapy. *Journal of the American Psychoanalytic Association, 11,* 722-733.

American Psychiatric Association (1994). *Diagnostic and Statistical Manual of Mental Disorders, 4th ed* (*DSM-V*). Washington, D.C.

Anderberg, M. (1973). *Cluster analysis for applications.* New York: Academic Press.

Antonovsky, A. (1987). *Unraveling the mystery of health.* San Francisco: Jossey-Bass.

Bachrach, H. M., Galatzer-Levy, R., Skolnikoff, A., & Waldron, Jr., S. (1991). On the efficacy of psychoanalysis. *Journal of the American Psychoanalytic Association, 39,* 871-916.

Beutler, L. E., Machado, P. P., & Neufeldt, S. A. (1994). Therapist variables. In Bergin, A. E., & Garfield, S. L (Eds.) *Handbook of Psychotherapy and Behavior Change*. New York: Wiley.

Bibring, E. (1954). Psychoanalysis and the dynamic psychotherapies. *Journal of the American Psychoanalytic Association, 2,* 745-770.

Blomberg J., Sandell R., Lazar A., & Schubert J. (1997, June). *Stability of effects after psychotherapy and psychoanalysis in terms of relapse-rate in the STOPP-study*. In: Abstracts of the 28th Annual Meeting, Society for Psychotherapy Research, June 25 - 29, 1997, Geilo, Norway (p. 67).

Bridges, K. & Goldberg, D. (1989). Self-administered scales of neurotic symptoms. In Thompson C (ed): *The Instruments of Psychiatric Research*. New York: Wiley, pp 157-176.

Carlsson, A-M. (1991). Privat psykoterapi finansierad med socialbidrag och bidrag från landstinget: En kartläggning inom Stockholms läns landsting. (Psychotherapy in private practice financed with social security or subsidy from the County Council: Mapping within the Stockholm County Council.) *Enheten för Psykosocial Forskning och Utveckling, Rapport 1991:*2, Nacka.

Carlsson, A-M. (1993). Psykoterapi finansierad med offentliga medel inom Stockholms läns landsting. Kartläggning av organisationen och en utvärdering. (Psychotherapy financed with public subsidy within Stockholm County Council. The organisation and an evaluation of the outcome.) *Enheten för Psykosocial Forskning och Utveckling, Rapport 1993:*2, Nacka.

Carroll, L (1865). *Alice's Adventures in Wonderland. In the Complete Illustrated Works of Lewis Carroll (pp. 11-114*). London: Chancellor Press, 1982.

Chambless, D. L. & Hollon, S. D. (1998). Defining empirically supported therapies. *Journal of Clinical and Consulting Psychology, 66,* 7-18.

Derogatis, L. R., Lipman, R. S., Rickels, K., Uhlenhuth, E. H., & Covi, L. (1974). The Hopkins Symptom Checklist (HSCL): A self-report symptom inventory. *Behavioural Science, 19,* 1-15.

Doidge, N. (1997). Empirical evidence for the efficacy of psychoanalytic psychotherapies and psychoanalysis: An overview. *Psychoanalytic Inquiry, 17,* 102-150.

Doidge, N., Simon, Gillies, L. A., & Ruskin, R. (1994). Characteristics of psychoanalytic patients under a nationalized health plan: DSM-III-R diagnoses, previous treatment, and childhood trauma. *American Journal of Psychiatry, 151,* 586-590.

Dossman, R., Kutter, P., Heinzel, R., & Wurmser, L. (1997). The long-term benefits of intensive psychotherapy: A view from Germany. *Psychoanalytic Inquiry, (suppl.),* 74-86.

Fisher, S., & Greenberg, R. P. (1996). *Freud scientifically reappraised. Testing the theories and therapy.* New York: Wiley.

Fosshage, J. L. (1991). Beyond the basic rule. In A. Goldberg (Ed.), *The Evolution of Self Psychology* (Vol. 7, pp 64-71). Hillsdale, NJ: The Analytic Press.

Fosshage, J. L. (1997). Psychoanalysis and psychoanalytic psychotherapy: is there a meaningful distinction in the process? *Psychoanalytic Psychology, 14* (3), 409 – 425.

Fromm-Reichmann, F. (1954). Psychoanalytic and general dynamic conception of theory and therapy. Differences and similarities. *Journal of the American Psychoanalytic Association, 11,* 711-721.

Garfield, S. L. (1986). Research on client variables in psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of Psychotherapy and Behavioural Change (3rd ed,.* pp 213-256). New York: Wiley.

Garfield, S. L. (1994). Research on client variables in psychotherapy. In A. E. Bergin & S. L. Garfield (red.). *Handbook of psychotherapy Research and behavior change (4:th ed*.). New York: Wiley.

Gill, M. M. (1954). Psychoanalysis and exploratory frames of psychotherapy. *Journal of the American Psychoanalytic Association, 2,* 771-797.

Gill, M. M. (1984). Psychoanalysis and psychotherapy: A revision. *International Review of Psychoanalysis, 11*, 161 – 179.

Gill, M. M. (1988). Converting psychotherapy into psychoanalysis. *Contemporary Psychoanalysis, 24*, 262-274.

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 Counselling Psychology, 32*, 3-22.

Hjelle, L. A., & Ziegler, D. J. (1981). *Personality Theories: Basic Assumptions, Research, and Applications.* New York: McGraw-Hill.

Jöreskog, K. G., & Sörbom, P. (1986). *LISREL VI. Analysis of linear structural relationships by maximum likelihood, instrumental variables, and least square methods* (*4th ed*.). Mooresville, IN: Scientific Software.
Kadushin, C. (1969). *Why People go to Psychiatrists?* New York: Atherton.

Kantrowitz, J. L. (1997). A brief review of psychoanalytic outcome research. *Psychoanalytic Inquiry (Suppl), 17*, 87-101.

Kernberg, O. F. (1999). Psychoanalysis, psychoanalytic psychotherapy and supportive psychotherapy: Contemporary controversies. *International Journal of Psycho-Analysis, 30* (6), 1075-1091.

Kline, R. B. (1998). *Principles and practice of structural equation modeling.* New York: Guilford.]

Lubin, B., Hornstra, R. K., Lewis, R. V., Bechtel, B. S. (1973). Correlates of initial treatment assignment in a community mental health center. *Archives of General Psychiatry, 29,* 497-504.

Olfson, M. & Pincus, H. A. (1994 a). Psychotherapy in the United States, II: Patterns of utilization. *American Journal of Psychiatry, 149,* 1289-1294.

Olfson, M. & Pincus, H. A. (1994 b). Psychotherapy in the United States, I: Volume, costs, and user characteristics. *American Journal of Psychiatry, 149,* 1645-1653.

Rangell, L. (1954). Similarities and differences between psychoanalysis and dynamic psychotherapy. *Journal of the American Psychoanalytic Association, 2,* 734-744.

Sandell, R., Blomberg, J., & Lazar, A. (2000). *Time Matters. On Temporal Interactions in Long-Term Psychotherapies*. Submitted for publication.

Schubert, J. & Blomberg, J. (1994). Psykoterapipatienter i offentlig vård. En social, demografisk och diagnostisk undersökning. (Psychotherapy patients in public health care. A social, demographic and diagnostic study.) *Rapporter från PI, nr 7*. Stockholms läns landstings Psykoterapiinstitut.

Shapiro, D. A. & Shapiro, D. (1982). Meta-analysis of comparative therapy outcome studies: A replication and refinement. *Psychological Bulletin*, *92*, 581-604.

Smith, M. L., Glass, G. V. & Miller, T. I. (1980). *The Benefits of Psychotherapy*. Baltimore: Johns Hopkins University Press.

Taylor, L., Adelman, H., & Kayser-Boyd, N. (1986). The Origin Climate Questionnaire as a tool for studying psychotherapeutic process. *Journal of Child and Adolescent Psychotherapy, 3*, 10-16.

Vessey, J. & Howard, K. (1993). Who seeks psychotherapy? *Psychotherapy*, *30*, 546-553.

Wallerstein, R. S. (1989). The Psychotherapy research project of the Menninger Foundation: An overview. *Journal of Consulting and Clinical Psychology*, *57*, 195-205.

Weber, J., Solomon, M., & Bachrach, H. (1985). Characteristics of psychoanalytic clinic patients: Report of the Columbia psychoanalytic centre research project (I). *International Review of Psycho-Analysis, 12*, 13-26.

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

Appendix

Standardisation of the Therapist Attitude Scales (TASC)

For standardisation purposes, the items in the TASC subscales were factor analysed, using data from the national random sample of licensed therapists. In an exploratory series of analyses, using various extraction and rotation methods and various principles for determining the number of factors, we found for each of the three item sets a stable and interpretable solution when three factors were extracted and rotated for each section. The highest-loading items in each such factor were subsequently used for forming factor scales. Thus, for each of the three first components in each set, a score was computed as the mean across all high-loading items (after transforming scores on items in the section on basic assumptions to a five-point scale, as for the other scales). As high-loading items we considered those that had routinely turned up with loading higher than, or around, .5 in the exploratory series of analyses. For items in the basic assumptions section our criterion had to be set a bit lower, at .4, in order to have enough items left. The items constituting each of the nine scales are listed in the end of the Appendix.

> As a validation of the scales, we tested whether they differentiated as expected between therapists with different theoretical orientations and with different schools of training as the basis for their licensing. The results of these tests are exhibited in Figures 5 and 6. For reasons of readability, the relatively small groups of therapists with family or group training were excluded from Figure 6. As will be seen, neither Irrationality nor Pessimism had any significant relations with theoretical orientation, as rated by the therapists themselves. They were also less strongly associated with training varieties. Further, Kindness discriminated not as well among groups with different theoretical orientations (Figure 5), otherwise, however, the TASC was strongly discriminative in a pattern that is wholly consistent with theoretical suppositions.

In order further to study the variation among the treatment providers, we then performed a series of cluster analyses. As before, we used the national random sample of therapists and applied Anderberg's (1973) nearest neighbour method as a suitable technique with large samples. After an exploratory series of analyses, we finally settled for a four-cluster solution which meaningfully and significantly related to different theoretical orientations and different training sites. The cluster profiles are shown in Figure 7. Thus, we found one cluster (12%) where therapists with cognitive or cognitive-behavioural training were over-represented; one (27%) with a rather complementary profile across the self-rating factors, with an over-representation of persons with psychoanalytic training (but also including persons with regular psychotherapeutic training); and two with profiles

high on the self-rating scales where the cognitive and cognitive-behavioural therapists were high and the psychoanalysts were low (Adjustment, Kindness, and Supportiveness) and *also* high on those where the psychoanalysts were high and the cognitive and behavioural ones were low (Insight and Neutrality). They differed radically on Self-disclosure, however, one cluster (34%) being high (like the cognitive/cognitive-behavioural cluster) and one (27%) being low (like the psychoanalytic cluster). Different training sites, with local particularities, were over-represented in the two clusters, but we chose to consider both as eclectic in their attitudes, endorsing *both* behavioural and psychoanalytic values. However, the profile of the larger of the two eclectic clusters was generally closer to the cognitive-behavioural cluster, whereas the profile of the smaller eclectic cluster was closer to the psychoanalytic cluster.

Table 3. The Therapeutic Attitudes Scale.

Curative Factors (”What do you think contributes to long-term and stable therapeutic change?”)

Promoting adjustment (“Adjustment”)
1. Giving the patient concrete goals
2. Working for adjustment to prevailing social circumstances
3. Helping P avoid anxiety-provoking situations
4. T takes the initiative and is leading the sessions
5. Stimulating P to think about his problems in more positive ways
6. Working with P’s symptoms
7. Giving P concrete advice
8. Helping P control his/her feelings
9. Helping P avoid repeating his/her mistakes
Promoting insight (“Insight”)
10. Helping P understand that old reactions and relations are repeated with T
11. Helping P see the connections between his/her problems and his/her childhood
12. Supporting P to ponder, in the therapy, painful early experiences
13. Working with P’s defences
14. Bringing P’s sexuality to the fore

15. P has the opportunity to work with his/her dreams
16. Help P understand that old behaviours and relations are being repeated
17. Interpret P's body language
18. Working with P's childhood memories

Showing kindness ("Kindness ")

19. The therapist is warm and kind
20. The patient feels well liked by the therapist
21. Supporting and encouraging the patient
22. Consideration and good care-taking
23. Let the patient get things off his chest

Therapeutic style factors ("What are you like as a therapist?")

Neutral attitude ("Neutrality")

1. I do not answer personal questions from the patient
2. I keep my personal opinions and circumstances completely outside the therapy
3. I am more neutral than personal in therapy
4. I do not express my own feelings in the sessions
5. My verbal intervention are brief and concise
6. Keeping the therapeutic frame is an important instrument in my work

Supportive attitude ("Supportiveness")

7. I often put questions to the patient
8. It is important to convey hope
9. It is important to order and structure the material
10. I am rather active in sessions

Self-disclosing attitude ("Self-disclosure")

11. I always communicate the therapeutic goals to the patient in the beginning of a therapy
12. I always make the therapeutic goals explicit to myself during a therapy
13. I admit my own mistakes to the patient

Basic assumptions ("What are your general beliefs about the human mind and about psychotherapy?")

Rationality v. irrationality ("Irrationality")

1. Human behavior is governed ... by free will/by uncontrollable factors
2. By nature, man is ... rational/irrational
3. Human behavior is governed ... by external, objective factors/internal, subjective factors

Craft v. Art ("Art")

4. Psychotherapy may be described ... as a craft/as free creative work
5. Therapeutic work is governed ... by training/by personality
6. Psychotherapy may be described ... as a science/as a form or art
7. Psychotherapeutic work is governed by … systematic thinking/intuition

Optimism v. pessimism ("Pessimism")

8. The basic principles of human behavior may be understood ... completely/not at all
9. Humans can develop ... infinitely/not at all
10. Therapeutic work is governed by the fact ... that everything may be understood/that *not* everything may be understood

Figures 5 to 7. Relationship between TASC Scores and Self-Reported Theoretical Orientation (upper panel), Type of Training, (middle panel), and Cluster (lower panel).

The development of a psychoanalytically-orientated day hospital treatment for borderline personality disorder: theory, problems, and practice.

Dr Anthony W Bateman, Consultant Psychiatrist in Psychotherapy, Barnet, Enfield, and Haringey Mental Health Care Trust & Honorary Senior Lecturer Royal Free and University College Medical Schools [1]

[1] Address : The Halliwick Unit, St Ann's Hospital, St Ann's Road, London N15 3TH

In this paper, I will discuss some general problems of research into personality disorder (PD), outline some of the difficulties faced in organizing research into borderline personality disorder (BPD), and finally summarize the results of the first randomized controlled trial of outcome of treatment of BPD in Britain (Bateman & Fonagy, 1999; Bateman & Fonagy, 2001)

Borderline personality disorder (BPD) is common, affecting about 1% of the general population and up to 20% of psychiatric in-patients. Approximately 9% of patients eventually kill themselves (Frances, 1986). This fact alone means that finding effective treatment is urgent and yet despite over two decades of research, our knowledge of the disorder and its treatment remains limited. The complexity of the disorder, characterised by 'stable instability', baffles clinicians and researchers alike. Characteristics of impulsivity, self-destructiveness, constant efforts to avoid real or imagined abandonment, chronic dysphoria, sudden anger or boredom, transient psychotic episodes or cognitive distortions, and identity disturbance all mitigate against a smooth ride for researchers and clinicians. It is therefore not surprising that many continue to avoid borderline patients finding them frustrating and impossible. Despite these problems, our unit set out to investigate a psychoanalytically orientated treatment for borderline personality disorder using a randomized controlled design. It soon became apparent why there was only one other randomized controlled trial of BPD (Linehan, Armstrong, Suarez, *et al*, 1991)!

Problems of outcome research into personality disorder

There are few controlled trials of treatment for BPD. There are a number of reasons for this. Firstly there are problems of case identification and co-morbidity. Although enshrined in diagnostic classifications, concern about the validity of the diagnosis remains. An overlap with affective disorders exists but Gunderson and Phillips (Gunderson & Phillips, 1991) have concluded that the two disorders are not the same although affective instability is at the heart of BPD. Secondly, BPD is an heterogeneous condition and varies in severity. Until recently there was no measure of severity making it impossible to assess the level of morbidity in a sample of patients. Crude attempts to establish severity at the outset of treatment or entry into a trial such as counting acts of self-harm over the preceding weeks or months probably bears little, if any, relationship to severity. A patient who makes serious suicide attempts at infrequent intervals may be more seriously disturbed than a patient who frequently, albeit usually sporadically, takes overdoses. Finally, there is the difficulty of random assignment which has now become the gold standard against which treatments are tested.

Randomisation

Borderline patients do not take kindly to randomisation. Their search is for stability and certainty. Offering them referral into a research project in which their allocation appears to be dependent on the toss of a coin confronts them with uncertainty and makes them vulnerable to fears of rejection. Both randomisation in and randomisation out causes problems. Borderline patients at the severe end of the spectrum have usually had years of psychiatric treatment and psychotherapy. Each new offer of treatment is a moment of hope. For those accepted into treatment, early expectations may not be met. When confronted with the reality of hard therapeutic work, the result may be, at best, a feeling of demoralisation and, at worst, rage and aggression and refusal to participate in any further aspect of research. Randomisation out of the treatment into a control group can lead to refusal to co-operate. Yet the researcher needs patients who are randomised out of the treatment programme to agree to further interviews and to fill out questionnaires. This can become progressively difficult over time leading to a high attrition rate in a control group, distorting the cohort of patients. Some patients may even take pleasure in ensuring that researchers do not get the information they ask for at the time that it is needed leading to further sampling problems. Given the relatively small cell sizes of RCTs, attrition represents a serious threat to internal validity. PD patients tend to show relatively high attrition rates in treatment trials (Tyrer, Seivewright, Ferguson, *et al*, 1990) although this varies according to personality disorder diagnosis (Shea, Pilkonis, Beckham, *et al*, 1990) and treatment approach(Linehan, Armstrong, Suarez, *et al*, 1991); (Rosser, Birch, Bond, *et al*, 1987); (Bateman & Fonagy, 1999).

In addition to the sampling problems discussed above there are other difficulties. First, there is an accumulating literature on the importance of patient expectations for therapy outcome (Horowitz, Rosenberg & Bartholomew, 1993). Strict randomisation may lead to treatment allocations incongruent with patient expectation and this may be particularly problematic for patients whose lack of flexibility is almost a defining feature of their disorder (Bleiberg, 1994). Second, RCTs, with notable exceptions (Shapiro, Rees, Barkham, *et al*, 1995), do not randomise therapists to patients even though it is known that the personality, skills and training of the therapist have significant effects on outcome (Beutler, Machado & Neufeldt, 1994). This potential confound is likely to be even greater for psychotherapeutic treatments of PD given that interpersonal relationship problems are undoubtedly at the core of personality disturbance. Third, investigator allegiance (Robinson, Berman & Neimeyer, 1990) has been shown to strongly affect outcome and unbiased, blind evaluations are hard to achieve in

long term treatments. Fourth, comparison groups are difficult to identify for long term therapy trials. A no treatment comparison is ethically unacceptable in BPD (Basham, 1986), particularly as suicidality and self harming behaviour are common. On the other hand, as long-term therapy tends not to be routinely available, a treatment-as-usual control group may be valid and appropriate.

There is a trade-off between the internal validity (Cooke & Campbell, 1979) of well-controlled trials, which ensure that causal inferences may be appropriately drawn from experimental manipulations, and the external validity of naturalistic research designs which are limited in terms of causal inference but which generate findings more readily generalisable to everyday practice (Jensen, Hibbs & Pilkonis, 1996) (Hoagwood, Hibbs, Brent, *et al*, 1995). At the extreme end of naturalistic studies are survey reports, such as the consumer survey of psychotherapy (Seligman, 1995). Of course, what is sacrificed in surveys is information about the exact nature of the treatment offered and information from individuals who did not respond. Imposing strict controls, however, carries a cost in depicting psychotherapy in a far more organised and coherent form than is available in the real world of the clinic. No wonder, then, that clinic-based studies tend to regularly under-perform more strictly performed laboratory-based investigations (Weisz, Donenberg, Han, *et al*, 1995).

But most importantly, the absence of clear distinction between even manualised treatment interventions has contributed to the lack of progress on specifying particular therapeutic approaches BPD. Outcome evaluation is hampered by the lack of specificity in psychological approaches to therapy (Roth, Fonagy & Parry, 1996) and some have argued that the considerable overlap between psychotherapies compromises the possibility of reaching conclusions concerning relative effectiveness (Goldfried, 1995). The problem is, once again, particularly acute in the case of the long-term approaches used in treating personality disorder. With such patients, practitioners make complex choices in selecting interventions that take account of both behavioural and dynamic factors. In order to enhance specificity researchers have “manualised” treatments and developed measures to assess the extent to which therapists are able to follow protocols outlined therein. Three approaches to therapy with borderline personality disorder have so far been manualised. These include psychoanalytic psychotherapy (Clarkin, Kernberg & Yeomans, 1999) ialectical behaviour therapy (Linehan, 1993b), and object relations/interpersonal approaches (Dawson, 1988) (Marziali.E, 1989). The manual for Cognitive-Analytic therapy is as yet untested (Ryle, 1997).

Treatment Intervention

The modified individual psychoanalytic approach adopted by Kernberg (Clarkin, Kernberg & Yeomans, 1999) is based on clarification, confrontation, and interpretation within a developing transference relationship between patient and therapist. Initially there is a focus and clarification of self-destructive behaviours both within and without therapy sessions. Gradually aspects of the self that are split off from the patient's core identity are challenged, especially as they impinge on chaotic impulsive behaviour, fluctuating affects and identity conflict which itself leads to dissociation. Understanding and resolving their impact on the transference relationship becomes central. Considerable work on elaborating and validating this therapeutic approach has been performed as part of an NIMH funded treatment development project demonstrating that it is possible to train clinicians to use this method (Clarkin, Foelsch, Levy, *et al*, 2001).

In contrast, Linehan's strategy in DBT uses support, social skills, education, contingency management, and alternative problem solving strategies to manage impulsive behaviour and affect dysregulation. A mix of both individual and group psychotherapy is used. However, the relationship between the patient and therapist is pivotal in helping the patient replace maladaptive actions such as self-destructive acts with adaptive responses during crises. Linehan (Linehan, 1993a) suggests that a number of aspects 'set if off from "usual" cognitive and behavioural therapy' and that 'the emphasis in DBT on therapy-interfering behaviours is more similar to the psychodynamic emphasis on "transference" behaviours than it is to any aspect of standard cognitive-behavioural therapies'.

The treatment strategy developed by Dawson (Dawson, 1988) and colleagues is named 'relationship management psychotherapy' (RMP). In essence, this approach conceptualises the borderline patient as struggling with conflicting aspects of the self, leading to instability. Interpersonal relationships, including the therapeutic relationship, become the context in which the patient tries to resolve conflicts through externalisation. For example, if a therapist is optimistic and active the patient becomes pessimistic and compliant. In some ways such polarities are similar to the reciprocal roles identified in cognitive analytic therapy. The task of the therapist's role is to alter the rigidity of the dialogue and to disconfirm the patient's distorted experience through attention to the process of sessions rather than the content of the interaction. The format is exclusively through time-limited group psychotherapy.

At first sight these three methods may sound distinctly different, ranging from individual therapy to a mix of individual and group therapy to solely group psychotherapy. Beyond that there are some striking similarities. Both Kernberg and Linehan focus initial sessions on the establishment and negotiation of a treatment contract within the framework of their approach. A particular emphasis is placed on self-destructive behaviour, especially therapy interfering behaviour, and appropriate limits are set and renewable contracts made. Both methods carefully define the responsibilities of the therapist on how self-destructive behaviour will be handled, regular appointments are arranged, and acceptance of difficulties of remaining in treatment are recognised and explicit statements made about the possibility of failure of treatment. Identity issues are central from a psychoanalytic viewpoint and therapists are constantly on the alert for split-off aspects of patients and how these are played out in the patient-therapist relationship. In DBT there is less emphasis on identity issues but nevertheless a 'black-and-white' cognitive style is targeted through dialectical techniques to help the patient overcome the all-or-none thinking and polarised approach to life. Both treatments prescribe the level of contact permissable between patient and therapist. In DBT, emergency sessions are allowed to enable the therapist and patient to develop alternative ways of crisis resolution other than hospital admission or self-destructive behaviour. In psychoanalytic therapy contact between sessions is not permitted although discussion of alternative routes to support between sessions may be a focus of a consultation. Implementation of the two treatments is consistent with theoretical views. Linehan provides information about cognitive-behavioural conceptualisation of self-destructive behaviour whilst Kernberg uses exploratory interpretations using ideographic hypotheses relating self-destructive behaviours to feelings about treatment. Both discuss alternative pathways to resolution of conflict and distress.

In contrast to these overlaps, RMP takes a more neutral stance. No formal contract is made, no attempt is made to interpret or to explain the patient's anger or self-destructive behaviour, and no emphasis is given to education or understanding about actions or threats that may disrupt therapy. Instead, the primary therapeutic task is to identify 'core messages' that reflect the polarities of conflict about which the patient is struggling. Therapists generate hypotheses about these as they are played out in the group setting whilst avoiding enacting any of the externalised, polarised selves. On theoretical grounds, it may be supposed that this is the least supportive therapy for borderline patients and

likely to lead to early drop out or failure to take up offer of treatment whilst DBT is the most supportive given its methods and the availability of the therapist. Whilst there is no data on the drop-out rate for RMP, Linehan has shown that the drop out rate is low in DBT (16%). Whilst the drop out rate for psychoanalytic therapy is reportedly higher, it may be altered. We (Bateman & Fonagy, 1999) had an attrition rate of only 12%.

The marked overlap between therapies for long-term treatment of personality disorder has significant implications for research since randomised comparison of one intervention with another sits uppermost in a hierarchy of stringent tests for any treatment. Not only may this control for many processes independent of the treatment and common to all psychological treatments but also may include tests between specific competing mechanisms. But 'horse-race' comparative studies in long term treatment are unlikely to be helpful in identifying better methods of treatment since there is so much variance within each treatment and overlap between them that differential treatment effects are likely to be masked. It is more important to isolate the effective aspects of different treatments (Waldinger & Gunderson, 1984). It remains unclear a decade on what the effective components of DBT are. The original and unreplicated study was from a university department using highly trained and supervised therapists with enthusiasm and motivation implementing a new approach. Whether this can be generalised using less trained personnel working in community teams remains to be seen. Davidson and Tyrer (Davidson, 1996) remark that the translation of such a complex treatment into limited resource settings such as community mental health centres is questionable given the many therapist hours and requirement for expert supervision. Of course, this issue is not peculiar to DBT and psychoanalytic treatments are probably poorly generalisable.

Conclusions from research

In the light of the considerable problems which still exist in conceptualising and defining borderline personality disorder, separating it from other mental disorders, and designing treatment trials of long term therapy (which have adequate internal as well as external validity), it is perhaps not surprising that our knowledge concerning effective psychological treatments of BPD seems still to be somewhat rudimentary. Effective treatment protocols are relatively few in number and even where they exist remain largely untested. However, studies consistently demonstrate modest gains associated with relatively high doses of treatment. There is also encouraging evidence that these gains are cost effective

(Gabbard, Lazar, Hornberger, *et al*, 1997) (Stevenson, 1999), particularly in terms of savings in health care costs.

Halliwick Psychotherapy Unit – research and its implications

Our research demonstrates many of the problems that are outlined above. Both its strengths and weaknesses arise from the fact that it is clinical service research resulting in a trade-off between internal and external validity. On the positive side, firstly the programme was developed and implemented by a team of generically-trained mental health professionals with an interest in psychoanalytically orientated psychiatry rather than by highly trained personnel within a university research department. Secondly, the research took place within a normal clinical setting and in a locality and healthcare system in which patients were unlikely to be able to obtain treatment elsewhere. The latter allowed effective tracing of patients within the service and accurate collection of data about psychiatric hospital admission. Thirdly patients were treated at only two local hospitals for medical emergencies such as self-harm, enabling us to obtain highly accurate data of episodes of self-harm requiring medical intervention. On the negative side, the programme was complex, leading to difficulty in being able to identify the effective ingredients should this be the result. It was also unfunded. However, the programme was designed so that it could be dismantled at a later date to determine the potent ingredients. At present a randomised controlled trial is underway of an out-patient treatment package made up of three of the ingredients that we consider to be the effective components of the programme.

In developing the research programme, we were joined by Peter Fonagy whose theoretical ideas and knowledge of research were pivotal in identifying a coherent treatment programme (Fonagy, Kennedy, Leigh, *et al*, 1992). Our initial tasks were to review the literature, to consider the evidence for effective interventions, and to match those to the skills within the team. We concluded that treatments shown to be effective with BPD had certain common features. They tended (a) to be well-structured, (b) to devote considerable effort to the enhancing of compliance, (c) to be clearly focussed, whether that focus was a problem behaviour such as self-harm or an aspect of interpersonal relationship patterns, (d) to be theoretically highly coherent to both therapist and patient, sometimes deliberately omitting information incompatible with the theory, (e) to be relatively long term, (f) to encourage a powerful attachment relationship between therapist and patient, enabling the therapist to adopt a relatively active

rather than a passive stance, and (g) to be well integrated with other services available to the patient. While some of these features may be those of a successful research study rather than those of a successful therapy, we concluded that the manner in which treatment protocols were constructed and delivered was probably as important in the success of treatment as the theoretically-driven interventions.

With these general features in mind, we set about developing a programme of treatment and organising a research programme to test the effectiveness of the intervention. From the outset it was clear that this was to be 'effectiveness research' rather than 'efficacy' research – we would investigate the outcome of BPD treated by generically-trained but non-specialist practitioners within a normal clinical setting. In this way, the treatment was more likely to be translatable to other NHS services without extensive and expensive additional training of personnel. But first we had to define a psychoanalytic view that was understandable to both staff and patients, second ensure that this enabled staff to think about any clinical situation that might arise, and finally, define how to react in a consistent manner to common situations such as suicide threats and acts of self-harm.

Psychotherapy , BPD, attachment and mentalizing

Psychotherapy, in all its incarnations, is about the rekindling of mentalization. Whether we look at Marsha Linehan's dialectic behaviour therapy protocol, John Clarkin's and Otto Kernberg's recommendations for psychoanalytic psychotherapy, or Anthony Ryle's cognitive analytic therapy, they all: (1) Aim to establish an attachment relationship with the patient, (2) Aim to use this to create an interpersonal context where understanding of mental states becomes a focus; (3) Attempt (mostly implicitly) to recreate a situation where the self is recognized as intentional and real by the therapist and this recognition is clearly perceived by the patient.

The core of our treatment programme for BPD is to a) help the patient understand and label emotional states with a view to strengthening the secondary representational system; b) enhance reflective processes; c) to focus on brief, specific interpretation, initially avoiding a focus on aggression. Enhancement of reflective processes enables the development of stable internal representations and the formation of a coherent sense of self. Care about interpretation is important. For example, the inevitable destructiveness of these patients in

relation to the therapeutic enterprise is rarely adequately dealt with by confrontation or interpretations of their aggressive intent. Such attacks are best regarded as self-protective.

Gaps in mentalization in BPD engender impulsivity and, during treatment, the intensification of the therapeutic relationship highlights the patient's difficulties and further exposes the rift between internal and external reality. This stimulates enactments. Attempts to bridge this dissociated mode of a patient's functioning, where nothing feels real (certainly not words or ideas) to moments when words and ideas carry unbelievable potency and destructiveness, can seem an awesome task. The therapist's concern is in some way analogous to that of the parents who create a frame for pretend play -- except in this case it is thoughts and feelings that need to become accessible through the creation of a transitional area. The therapist must get used to working with precursors of mentalization. The task is the elaboration of teleological models into intentional ones (Dennett, 1987). Yet it is only by being able to become part of the patient's pretend world, trying to make it real, while at the same time avoiding entanglement with the equation of thoughts and reality, that progress becomes conceivable. In our view, this process is best done within a transference-countertransference relationship but by a team of professionals rather than by an individual working alone because of the severe difficulty in avoiding destructive entanglements.

Transference

Whatever the approach taken to the treatment of BPD, problems of transference and countertransference are inevitably present and need to be planned for. Even in DBT, supervision takes into account the feelings engendered in the therapist by that patient and how such feelings can distract the therapist from his task. But should the psychoanalytic therapist work in the transference with borderline patients?

The transference of early relationship patterns onto current relationships, while ever present, is rarely helpful to highlight. Without mentalization, which acts as a buffer between internal and external reality, transference is not displacement but is experienced as real. If the therapist is experienced as an abuser he is the abuser -- no "as-if" about it. When such transference interpretations are made, the patient is often thrown into confusion and to protect the therapy has no choice but to enter a pretend mode in which their subjective experience has no relationship to what is perceived by the therapist as reality (Fonagy & Target, 1996). Gradually patient and therapist may elaborate a world, which however

detailed and complex, has little experiential contact with anything that feels real. In our view transference interpretation has to be more circumspect and is best dismantled into small parts that build up over time in an incremental way. For example, a simple acknowledgement of affect in the here and now, while conveying in words, tone and posture that the therapist is able to cope with the patient's emotional state may be the most productive line initially. Generic transference interpretation should only be used, if ever in its raw form, only late in treatment. Transference, using the term in its broadest sense, is helpful as a concrete demonstration of alternative perspectives. The contrast between the patient's perception of the therapist as she or he is imagined and as she or he actually is may help to place quotation marks around the transference experience.

Some programmes attempt to control enactments by making therapy contractually dependent. In our day hospital and out-patient programme we do not make 'therapy dependent' contracts. To do so risks discharging the patient for the very problems for which they are being treated. Being modest in one's aims is the most helpful device. One should not hope that insight through interpretation of transeference will prevent enactment; the aim is simply the gradual encouragement of mentalization. Consequently, the interpretation of enactments is rarely as helpful as trying to deal with their antecedents and consequences. We need to be equally permissive about our own tendency to enact in the countertransference. We have to accept that in order for the patient to stay in mental proximity we have to become what they need us to be. Yet we know that if we become that person, we can be of no help to them. Our aim should be the achievement of a state of equipoise between the two - allowing oneself to do as required yet trying to retain in our mind as clear and as coherent an image of the state of the mind of the patient as we are able to achieve.

Split transference

One of the most complicated challenges arising from treating BPD relates to externalisations of unbearable self-states. Splitting the transference by creating alternative foci for the patient's feelings is important here. In our programme the transference is split in a number of ways. Firstly, a package of group and individual therapy splits the transference and allows the patient to reflect on himself in the group during the individual session. Secondly, patients with BPD commonly have severe social problems or trouble with the law and so an additional member of the team is appointed to help them deal with these practical

realities whilst the individual therapist focuses on the relationship problems, unencumbered by practical issues.

So, what are the hallmarks of a successful therapy with an individual with severe borderline features?

No theory gets anywhere close to explaining the complex problems of this group of patients. However, having a theoretically coherent approach is vital. Such patients require that we are predictable and our implicit working models of them can then begin to form the core of their self-representations. A stable, coherent image is impossible to maintain, should the therapist swap theoretical approaches at an alarming rate. Mentalization can only be acquired in the context of an attachment relationship. This means that the therapy must embody a secure base. Attachment is inseparable from a focus on the mental state of the other. There can be no bond without understanding, even if understanding is possible without a bond. Treatments always take considerable time, and consistency over such prolonged periods is often hard to maintain. The patient is terrified of and actively fights mental closeness, even when physical proximity appears to be his overarching goal. Retaining such proximity while under persistent attack is neither comfortable nor likely to be achieved unless one leaves one's personal sensitivity at the door. Finally, one should be careful not to under-estimate the extent of the patient's incapacity. It is easy and relatively comforting to engage with the representational world of these patients at a level of complexity that they, in reality, have little appreciation of. They are readily seduced into such relationships and accept these complexities within a pretend mode, dramatically removed from anything which feels real to them. Such therapies tend to be durable but they are sadly unhelpful in the long run.

In order to establish consistency within a secure base and to minimise entanglement within transference and countertransference enactments we take a team approach. The team's mentalistic, elaborative stance ultimately enables the patient to find himself in the team's mind as a thinking, feeling being. This allows him to integrate this image as part of his sense of himself. There is a gradual transformation of a non-reflective mode of experiencing the internal world which forces the equation of the internal and external into one in which the internal world is treated with more circumspection and respect and as separate and qualitatively different from physical reality. Even if work were to stop here, much would have been achieved in terms of making behaviour understandable, meaningful and predictable. The internalisation of the team's

concern with mental states enhances the patient's capacity for similar concern towards his own experience. Respect for minds generates respect for self, respect for other and ultimately respect for the human community. It is this respect which drives and organizes the therapeutic endeavour within our programme and it is the operationalisation of these ideas that we put to the test.

Research and Results

In the present study, we carried out a randomisation of patients either to treatment in the day hospital programme or to continuing treatment within the general psychiatric service (control group). All patients were assessed using standardised criteria for borderline personality disorder, namely the Structured Clinical Interview for the DSM-III-R (SCID-II) (Spitzer, Williams, Gibbon, *et al*, 1991) and the Diagnostic Interview for Borderlines (DIB) (Gunderson, Kolb & Austin, 1981). A cut off score of 7 or more was used for a formal diagnosis of BPD. If patients met both criteria for BPD they were selected for randomisation either to treatment in the day hospital programme or to continuing psychiatric treatment. Patients were excluded from the study if they also met DSM-III-R, based on SCID-I criteria, for schizophrenia, bipolar disorder, substance misuse, or mental impairment, or had evidence of organic brain disorder based on SCID-I (Spitzer, Williams, Gibbon, *et al*, 1990). 60 referrals met the criteria for inclusion in the study. 10 refused to participate in the randomisation. 6 of these were admitted to the day hospital programme and excluded from the present study and 4 declined further treatment of any type. 6 further patients did not wish to participate in regular self-assessment and so were also not included. This process of randomisation sounds easy but in fact borderline patients change their mind about research on a regular basis and it becomes increasingly difficult to ensure that patients are clear about their decisions. However, when everything was sorted out there were no significant differences on any of the baseline measures for patients who did not participate in the study compared with those that entered the study. This left 44 patients entering the study who were randomly assigned to the two groups. Within the first month of entering the study, 3 control patients crossed over into the day hospital programme following serious suicide attempts leading to in-patient medical and psychiatric treatment. 3 patients (12%) in the day hospital group dropped out of treatment within 6 months. All were available for follow-up. No subjects dropped out of the control group. Demographic and clinical characteristics of the total cohort of patients are described in the original paper. Following randomisation there was no significant difference on any variable between the two groups including frequencies or average number of axis 1 and

axis 2 disorders. Particularly notable was the association of mood and anxiety disorders with BPD.

Treatment in the day hospital condition consisted of: (1) once-weekly individual psychoanalytic psychotherapy, (2) three times per week group analytic psychotherapy lasting an hour each, (3) once a week expressive therapy informed by psychodrama techniques (1 hour), (4) weekly community meeting (1 hour), all spread over 5 days; in addition, on a once per month basis, subjects had (5) a meeting with the case-administrator (1 hour), and (6) medication review by the resident psychiatrist. Therapies and informal patient-staff contact were organised in accordance with a psychoanalytic model of BPD as described above. Medication consisted of antidepressant and anti-psychotic drugs prescribed as appropriate, polypharmacy was discouraged. The maximum length of treatment was set at 18 months.

All therapy was given by psychiatrically trained nurse members of the day hospital team with no formal psychotherapy qualifications. Adherence to therapy was monitored through supervision (twice per week with the whole team) using verbatim session reports and by completion of a monitoring form collecting information about activities and interventions of therapists. Aspects of the day hospital programme have been described elsewhere (Bateman, 1995; Bateman, 1997).

We chose ‘treatment as usual’ in the general psychiatric service as control treatment. This consisted of (1) regular psychiatric review with the senior psychiatrist when necessary (on average twice per month), (2) in-patient admission as appropriate (admission rate 90%, average duration 11.6 days) with discharge to non-psychoanalytic psychiatric day hospital treatment focussing on problem solving (72% attended day hospital with average length of stay of 6 months), followed by (3) outpatient and community follow-up (100%, fortnightly by CPN visits) as standard aftercare. None of the control group received any formal psychotherapy. The initial types and doses of medication were the same for both groups. While this group cannot be considered to have received comparable amount of professional attention to the day hospital group, the approach controls for spontaneous remission.

Measures of outcome

Although we used a series of self-report measures, only the effectiveness of the programme in reducing suicide attempts and other acts of self-harm, decreasing hospital admissions, and ameliorating depression will be considered here.

a) Acts of self-harm and Clinical measures

The criteria for suicidal acts were: 1) deliberate; 2) life threatening; 3) had resulted in medical intervention; 4) medical assessment was consistent with a suicide attempt. Criteria for acts of self-mutilation were: 1) deliberate; 2) resulting in visible tissue damage; 3) nursing or medical intervention required.

A semi-structured interview (Suicide and Self-harm Inventory) was used to obtain details of both suicidal and self-damaging acts for the 6 month period before patients entered the study. This interview asks specific questions not only about numbers of acts but also about dangerousness of acts, i.e. presence or absence of another person, likelihood of being found, preparation, and lethality. Multiple acts over a short period of time, for example a frenzied self-cutting, were counted as a single act. Day hospital patients were monitored carefully with regard to self-destructive acts and control patients were interviewed every 6 months. Self-reports of suicidal and self-mutilatory acts were cross-checked with medical and psychiatric notes.

For all patients, a search of the hospital in-patient database was made to obtain the number of hospital admissions and the length of stay during a period of 6 months before entry into the study. This was cross-checked with the medical notes. All patients were admitted to the local unit because of the contracted nature of the service. Hospital admission and length of stay and psychiatric day hospital programme attendance was monitored throughout the study for all patients.

Follow-up

An attempt was made at 18 months following admission to follow all 44 patients for an additional 18 months. No patient in the partial hospitalization program was lost to follow-up, but some refused to complete all assessments at all time points. Three patients in the control group refused continued participation. Complete medical records were, however, available for these patients. While assessments were not blind, all the outcome variables were based on objective

clinical records confirmed by independent evaluation or were self-report measures.

Details were collected of both suicidal and self-damaging acts at the 24-, 30-, and 36-month evaluations. For all patients, searches of the hospital inpatient database were made at the 24-, 30-, and 36-month evaluations to obtain the number of hospital admissions and the lengths of stay over the preceding 6 months.

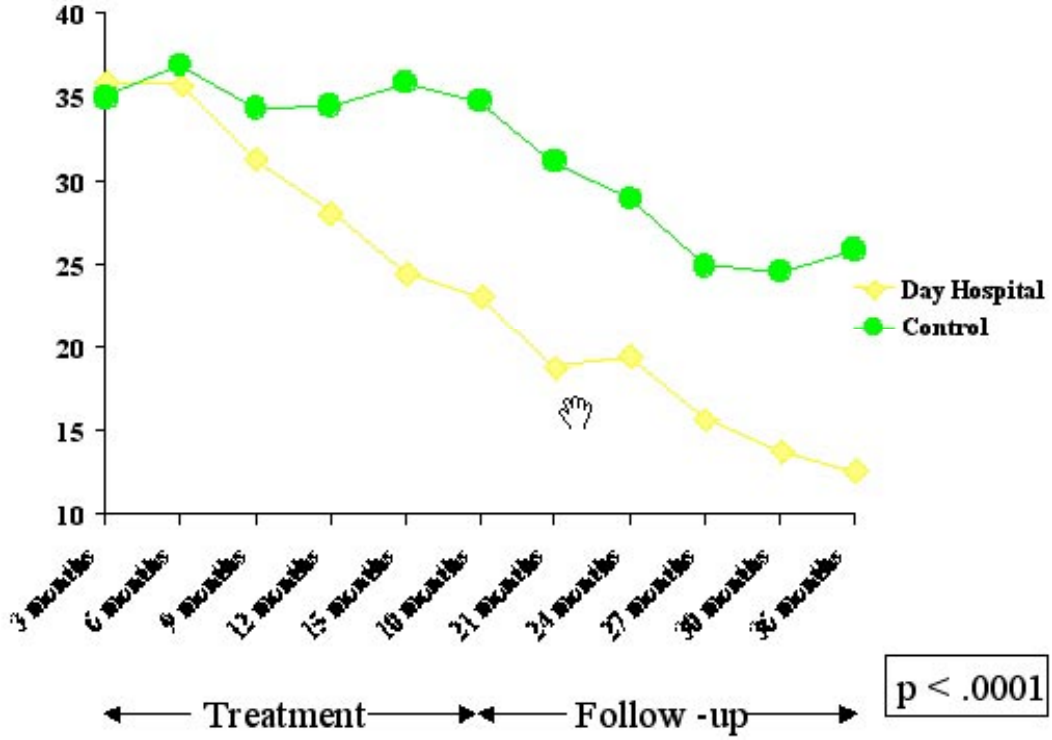
It was not possible to prevent patients having further treatment. Participation in other treatment programs was monitored throughout the study for all patients, including medication data ascertained from prescription charts and dispensing records. A follow-up program was offered to the patients assigned to the partial hospitalization program, which was attended by all except the three who prematurely terminated treatment. The program consisted of group analytic therapy twice a week (180 hours over 18 months) and review in a psychiatric outpatient clinic if requested every 3 months. Group attendance was 75% during the follow-up period, which indicates the stability of this cohort of treated patients. Community centre attendance and general psychiatric partial hospitalization programs were available through self-referral. The control group continued their general psychiatric treatment, which could involve inpatient admission when required, a general psychiatric partial hospitalization program, outpatient consultation, community centre attendance, or medication. None received any formal psychotherapy, although this was not precluded during the follow-up period.

Results

Detailed results can be found in our two published papers (Bateman, 1999; Bateman, 2001) and only some points of particular interest will be discussed here.

Depression (see fig 1) -

Figure 1 Self Rated Depression (Beck)



Our treatment programme made little difference to self-reported symptoms of depression as measured by the Beck Depression Inventory (Beck, 1961) for 6 months but following that period a continual decline in depressive symptoms was noted. At discharge only 3 treated subjects and no controls were below the clinical cut point on the BDI. The proportion scoring below 14 increased over the follow-up period to 59% by 18 months in the treated group but only 12.5% in the controls.

This is in contrast to the RCT of dialectical behaviour therapy in which there were no changes in levels of depression either at the end of treatment or during follow-up. It seems that the psychoanalytically-oriented programme stimulates rehabilitative effects but a cognitive behavioural programme focusing on symptoms and skills does not. This argument is further supported by the results of suicide and self-harm during follow-up.

Suicide and self-harm (see figures 2 and 3) - *Insert figures 2and 3*

Figure 2 % Attempted Suicide

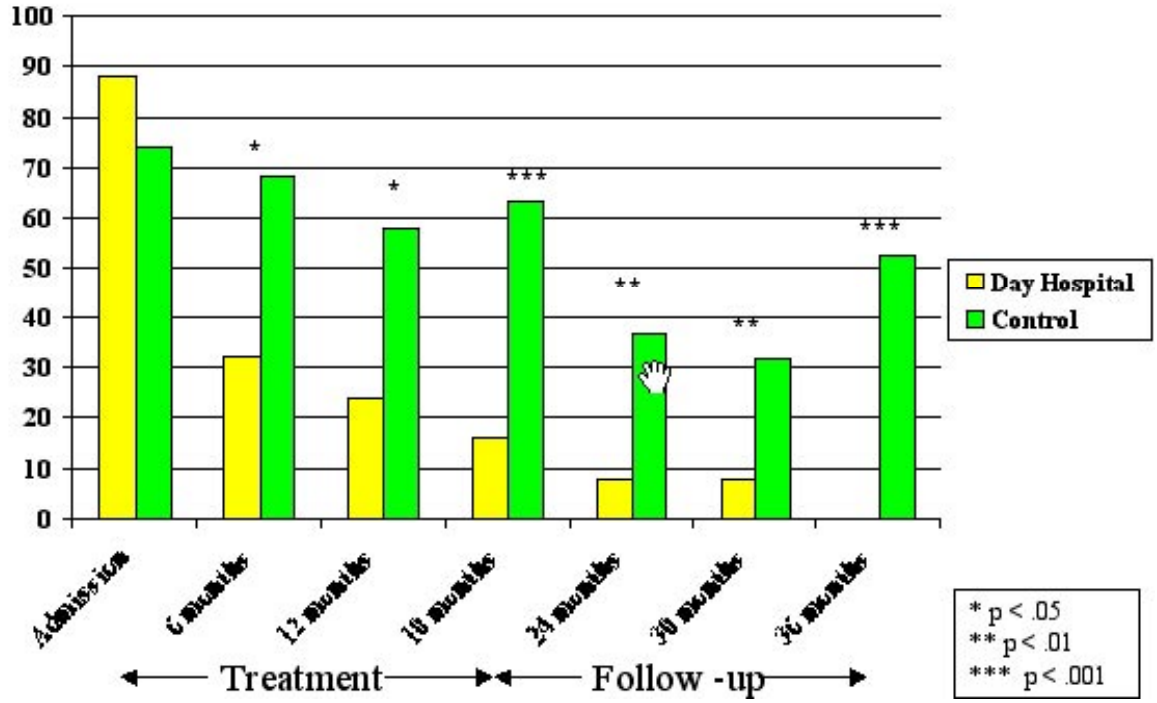
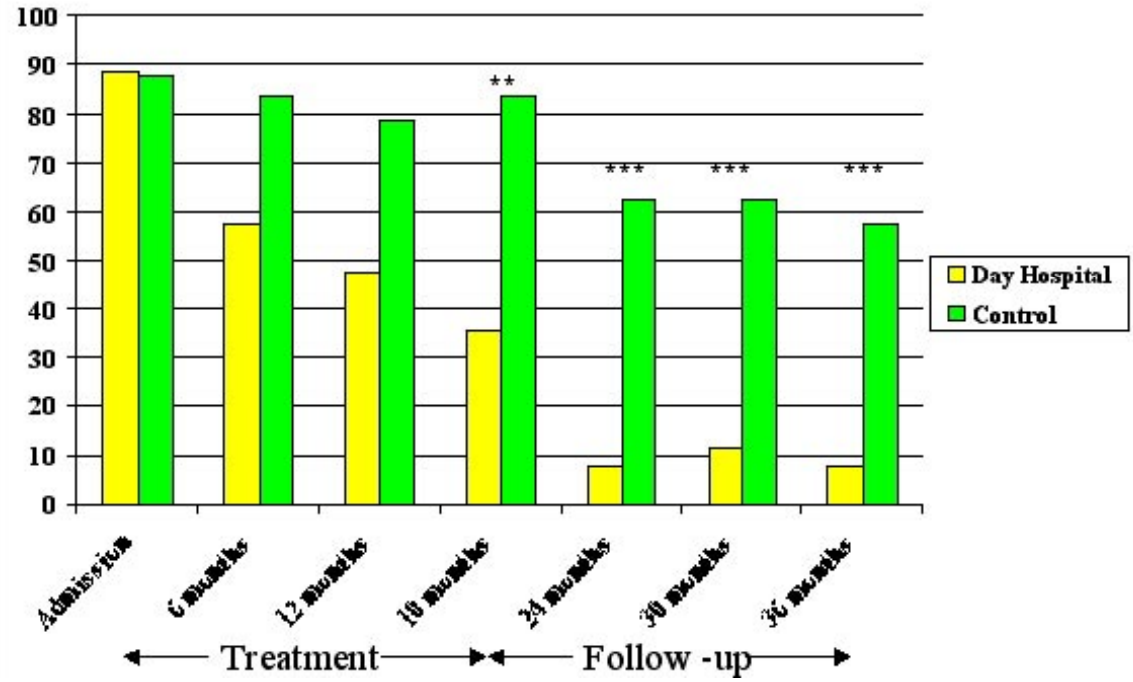


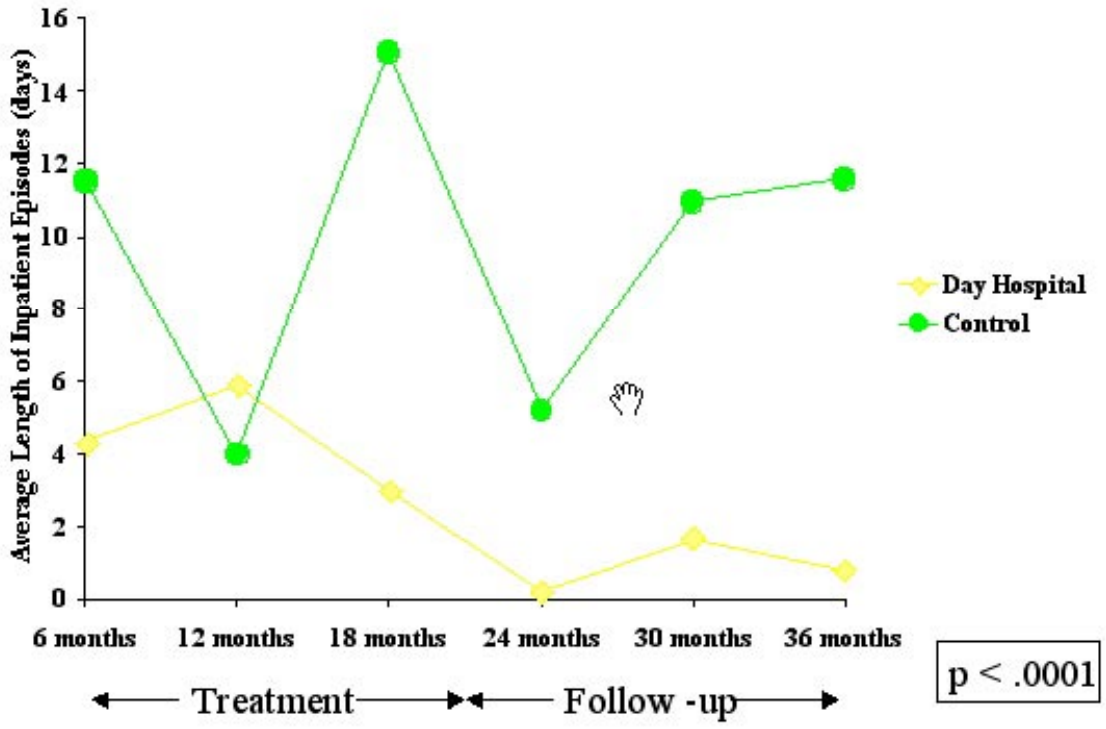
Figure 3 % Self-Mutilating Behavior



The continual decline throughout follow-up in suicide attempts and acts of self-harm in the treated group compared with the control group is testament to the rehabilitative effects of the programme. Throughout the study and follow-up period there were no successful suicides in either group and the rapid decrease in suicide attempts during the first 6 months of treatment suggest that simply offering patients a structured, coherent programme of treatment may suffice to reduce their dangerous behaviour. The slower effect on self-harm, with a significant change occurring after 1 year, suggests either that our programme did not focus adequately on such acts (it does not specifically target self-harm alone) or that understanding the meaning of self-laceration in terms of interpersonal and affective contexts takes time to have an effect on such symptoms as a way of dealing with anxiety.

Psychiatric admissions (see Figure 4) - Insert figure 4

Figure 4 Length of Inpatient Episodes



The average length of hospitalisation throughout treatment and follow-up, adjusted for pre-admission values, is displayed in Figure X. This confirms that average length of hospitalisation in the control group in the last six months of the study increased dramatically whilst in the PH group it remained relatively stable at around four days per six months. The group-by-time interaction was significant ($F=7.7$, $df=1,35$, $p<.01$), with a highly significant quadratic

component (F=13.3, df=1,35, p<.001). The post-hoc test yielded significant differences at 6 months (t=7.66, df=36, p<.001) and 18 months (t=13.23, df=36, p<.001). An identical pattern emerged for number of in-patient episodes (F = 14.1, df = 1,35, p<.001; F=19.9, df=1,35, p<.001 for the two-way and quadratic component of the interaction respectively). In the day hospital group no patient was admitted 6 months after discharge but 1 was admitted for 20 days 1 year after discharge and a further patient was admitted twice in the final 6 months of follow-up for 25 and 12 days respectively. In contrast, in the control group, 7 patients were admitted at least once during the first 6 months after discharge, 7 during the second 6 months, and 14 during the final 6 months. These differences were all significant on Fishers Exact Test (p<0.002, p<0.02, p<0.001 for 6, 12, 18 month time points of follow-up respectively). The average number of days in hospital increased from 6 (SD=10.8, range 0-28) in the first 6 months of the follow-up period, to 12.7 (SD=19.4, range 0-65) in the second, and 15.8 (SD=12.9, range 0-40) in the final 6 months of the follow-up. The differences were significant at all time points on the Mann-Whitney test (U=143, n=41, p<0.005; U=138, n=41, p<0.007; U=72, n=41, p<0.001 for 6,12, 18 months respectively). As there was little variation in the number of hospital days in the day hospital group we only examined trends in the control group. The repeated measures ANOVA on this group indicated significant differences between time points of assessment (Wilk's Lambda=.644, F=4.18, df=2,16, p<0.03). Exploring the polynomial components of this effect confirmed a significant quadratic effect (F=5.60, df=1,17, p<0.03) and no significant linear effect (F<1, df=1,17, NS).

Conclusions

In the treatment of BPD, effectiveness of treatment can only really be shown though prolonged follow-up. BPD is a relapsing and remitting problem with individuals showing periods of reasonable function followed by episodes of chaos and disorder. Only prolonged follow-up can determine if greater stability has occurred. Our study has one of the longest follow-up periods of a randomised controlled trial of treatment. An uncontrolled study of psychoanalytically orientated treatment by Stevenson and Meares (Stevenson & Meares, 1992; Stevenson & Meares, 1995) has data from a 5-year follow-up. Both follow-up studies show that the initial gains found at the end of treatment are maintained during follow-up. Howard’s interesting conceptualisation of psychotherapy into three phases is helpful in understanding this. He suggested that psychotherapy process could be understood as firstly showing a period of

remoralization, secondly remediation, and finally rehabilitation. Remoralization or a reduction in distress tends to occur quickly and some of the early changes in our patients such as the fall in general symptom distress were possibly a result of remoralization. Remediation involves refocusing the patients coping skills and helping them see their problems from a different perspective. This is, of course, a major aspect of our programme as we attempt to increase mentalisation and identification and self-control of affect. It is during this time that patients realise that their problems result from longstanding patterns that are maladaptive and unconsciously determined and that there are no quick fixes. But the real test of treatment is whether or not there are any rehabilitative effects of the programme and this is only determined by follow-up. Do patients who have made gains at the end of treatment maintain those gains? Can they cope with the everyday stresses and strains of life without engaging in their previous strategies of self-harm and so on? Our results suggest that this is the case. Not only are the gains made at the end of treatment maintained, but there are also further improvements. This is particularly clear in the results of depression, suicide attempts and self-harm, and psychiatric admissions.

Whilst the results of our study are encouraging, we neither know why patients improve, nor which patients are likely to benefit most from a psychoanalytically orientated treatment. Further development of psychoanalytic approaches to the treatment of BPD will only come about if we identify more precisely the mechanisms of therapeutic change and we decide on the sequencing of interventions and on whom the interventions are to be carried out. If psychoanalysis is to remain a vibrant and living discipline further research is urgently needed. Only if this takes place will a psychoanalytically based treatment of BPD have a central place in the 21st century.

References

EN.REFLIST

HENDERSON HOSPITAL DEMOCRATIC THERAPEUTIC COMMUNITY: OUTCOME STUDIES AND METHODOLOGICAL ISSUES

Fiona Warren and Kingsley Norton

INTRODUCTION

History of Henderson Hospital

Henderson Hospital is a democratic therapeutic community, based in South London. It was set up in 1947 by Dr. Maxwell Jones, one of the pioneers of therapeutic community developments within the UK. The aptly named "Industrial Neurosis Unit" was started on an experimental basis with the implicit assumption that its utility required evaluation. In establishing the unit, the then Ministry of Labour's aim was to help resettle "the industrial misfit" following World War II (Whiteley 1980). The idea behind the experiment was simple: if during the war, a useful place had been found for society's social and psychological casualties, so it should also be possible in peacetime for them to be helped to find resources within themselves to lead more useful and fulfilling lives.

Over time, the patient group, treatment methods and aims and objectives of the Unit changed, as did its name. It became the Social Rehabilitation Unit in 1954, in acknowledgement that the inability of many to settle in a job was due more to their lack of social skills than industrial or educational expertise. Patients were increasingly referred via the Courts, rehabilitation offices or social work agencies to which they had drifted and become labelled as social inadequates. By the mid-fifties the Unit was recognised as a centre specialising in therapeutic community ideology and the treatment of "psychopaths" (Whiteley 1980). In 1959, the Unit was again renamed, Henderson Hospital, when it became an independent of its large parent hospital – Belmont Hospital. Its naming honoured Professor D.K. Henderson, author of "Psychopathic States" (Henderson 1939).

Over the ensuing forty years, the Henderson Hospital approach has stabilised, maintaining an equilibrium which continues to address the therapeutic needs of its severely personality disordered clientele, despite periodic threats to its survival resulting from its highly specialist (and unusual) status. The treatment has been described in increasingly greater detail, enhancing the chances that the treatment approach could be replicated and the generalisability of it's outcome findings evaluated (Whiteley 1986; Norton 1992a; Norton 1992b; Norton 1995; Norton 1996) and more recently, the hospital has collaborated with two other UK Mental Health NHS trusts (South Birmingham and Mental Health Services of Salford) in an attempt to replicate the model.

Henderson Hospital today

The aim of much of the integrated socio- and psycho-therapy programme can be summarised as converting people who "act", rather than thinking or feeling, into individuals who are able to sustain and recognise their feelings and begin to articulate their psychological needs (Norton and Dolan 1995a). Maintaining this approach throughout 24 hours per day is problematic, entailing as it does staff working in close collaboration with the patients themselves (Norton and Hinshelwood 1996). Achieving this overall aim necessarily involves staff in an ongoing set of transference-countertransference relationships with patients, which spans (potentially) 24 hours of the day (Whiteley 1986). This psychodynamic situation requires an elaborate set of supportive and supervisory structures for staff in order for them to sustain, reflect on, and process their countertransference reactions (Rosenbluth 1991). With this kind of client group such activity is extremely important in minimising the destructive effects of any inter-staff splitting and promoting a therapeutic culture (Main 1957; Norton and Hinshelwood 1996).

Early literature on the TC suggested various definitions which may still be seen to underlie the approach today: 'a community with the immediate aim of full participation of all its members in its daily life with the eventual aim of re-socialisation of the neurotic individual for life in ordinary society' (Main 1946). For Maxwell Jones, the TC implied that: "the responsibility for treatment is not confined to the trained medical staff but is also the concern of other community members, i.e. the patients' (Jones 1952). Perhaps Main captured the democratic TC's essence with this definition:

This last definition also encapsulates the high degree of complexity within the treatment model and implies some of the difficulties with both its accurate description, or "manualization" for example, and its adequate evaluation (Norton 1992a).

The Past: Retrospective Outcome Studies

Social criteria of successful treatment: clinical improvement, employment, reconviction and readmission to psychiatric facilities

Clinical improvement

The Rapoport team, brought together to study the Belmont Social Rehabilitation Unit, intended to evaluate the treatment programme for which critics of Maxwell Jones and his method were calling (Rapoport 1960). However, there was controversy over the diagnostic use of the term "psychopath" and, partly because of this, the study changed direction, becoming more of an in-depth exploration of the workings of the community than an outcome study (Whiteley 1980). The

overall study period spanned four years (1953 to 1957) (Rapoport 1960; Manning and Rapoport 1976).

A small-scale outcome study was incorporated, however, and the findings suggested that the democratic therapeutic community method was not universally applicable, that some selection process for patients was necessary and that the intensive social and interpersonal pressures could damage those with weak ego structures. Rapoport had also pointed out a conflict of aims between those of the therapeutic staff and those of the workshop instructors who were aiming for "rehabilitation". The findings were largely viewed negatively by the Unit's staff and Maxwell Jones himself left the Unit in 1959 (Whiteley 1980).

For the outcome study Rapoport's team personally interviewed and classified 64 patients, one year after discharge, according to whether they were 'improved', 'same' or 'worse' compared to when they entered the Unit. Forty-one per cent were considered improved, eighteen per cent unchanged and thirty-one per cent worse, on the stated criteria. Improvement was associated with longer duration in treatment, fifty-two per cent for those staying more than 200 days were judged improved. No objective measures were used nor were any comparison samples studied.

Reoffending, employment and responsibility

In a later outcome study, a postal survey of the probation officers of eighty-six consecutive male discharges, on probation or borstal licences, sixty-two per cent of those traced were free of further convictions, up to twenty-two months post-discharge (Tuxford 1961). The response rate of the study was eighty-four per cent. Assessment was made using a four-point scale completed by the probation officer: from 1 (-increased sense of responsibility, employment and no further offending) to 4 (-further offending, unemployment and lack of responsibility). Twenty-four per cent fell into the first category, thirty-one per cent into the second, twenty-eight per cent were in the poor outcome category and seventeen per cent considered as complete treatment failures (Dolan and Coid 1993). Overall this represented a fifty-five per cent success rate, on the assessment criteria.

A similar rate of success was found in a follow-up study of discharged men (number unknown) who were assessed at 9 months post-discharge (Taylor 1963). Twenty-two per cent had found their own employment and a further forty-seven per cent had been placed in employment with professional support. The latter

group was followed up for a further 9 months and sixty per cent of these were still in work with a satisfactory report from their employers.

Forty-five patients had no further psychiatric admission or conviction at 2- year follow-up. Of men with previous convictions, forty per cent remained free of conviction. Of those with previous admissions to a psychiatric hospital, almost sixty per cent remained out of hospital over the two-year period. Good prognostic factors were: some level of school achievement; ever sustaining employment for more than two years; higher social class occupation; ever having been married and a history of affective disorder. A negative outcome was associated with: having previous convictions; a prior prison sentence(s); a probation order at referral or admission; current court proceedings; and institutionalisation before the age of 15.

A further outcome study was undertaken to develop a prediction equation for successful outcome (Copas and Whiteley 1976). Two cohorts of male patients were studied. One cohort of 104, at two years, showed forty-two per cent as having no further convictions or re-admissions, while the cohort of eighty-seven revealed a slightly higher figure of forty-seven per cent successful outcome on these criteria. At 5 years follow-up of the 104, one third had no reconviction or readmission. A further eleven per cent had only minor ‘relapse’ in the first year of follow-up, remaining free of conviction or relapse in the succeeding four years. It could be considered therefore that forty-five per cent of the total had a good outcome.

None of these early outcome studies utilized any comparison groups making it impossible to conclude that improvements were attributable to Henderson Hospital’s democratic therapeutic community. However, a 5-year follow-up study of 194 (male and female) patients was carried out which also reported on fifty-one patients referred to Henderson but not admitted (Copas, O’Brien et al. 1984). Similar criteria of success were utilised. At 3 years, forty-one per cent of the treated sample was improved compared to twenty-three per cent of the non-admitted. At 5 years, the relative proportions were thirty-six and nineteen per cent. There was no significant gender difference. Further analysis of these findings showed a positive association between success rate and length of stay in treatment. At 3 years’ follow-up, sixty-two per cent of those who stayed for 6 months and seventy-one per cent of those staying for 9 months, were improved. At 5 years, the relevant proportions were fifty-seven per cent and sixty-five per

cent for 6 and 9 months' stays, compared with nineteen per cent of those not admitted (i.e. having treatment as usual elsewhere).

Psychological criteria: neurotic symptomatology

Reconviction and readmission to psychiatric facilities are often used as outcomes but are only indirect measures of psychological health. These earlier studies of outcome (measured mainly in terms of further conviction or psychiatric hospitalisation), were later complemented by the undertaking of a study of psychological morbidity (Dolan et al. 1992), as measured by the Symptom Check-list (SCL-90-R; Derogatis, Rickels et al. 1976). Sixty-two subjects were followed up at an average of 8 months post-discharge and findings revealed a highly significant improvement on the Global Symptom Index (i.e. total score on SCL-90-R) post-discharge. Again, this was not a controlled study and the numbers were relatively small, although representing a typical figure for the number of patients treated in a given year.

The SCL-90-R is a self-rated instrument. However, data were subjected to a rigorous statistical analysis, addressing both the issue of reliability and also the importance of the observed clinical change (Jacobson and Truax 1991). With this method, fifty-five per cent of the group had improved reliably (i.e. had moved two standard deviations in relation to their baseline score). In thirty-two per cent, the change was clinically significant (i.e. subjects no longer scored in the pathological range but had scores within the normal range for the measure, defined by reference to normative data published on the instrument). Only 6.5 per cent had deteriorated reliably.

In this sample, length of stay was not significantly related to outcome, although those who stayed longer than nine months to show greater or beneficial change than those who stayed less (mean improvement .73, s.d. .84 for the longer stay vs .58, .86, for he medium stay and .61, 1.1 for the short stay groups). However, gender was significantly related to length of stay, more women staying longer than nine months (23/44, 52% vs 12/51, 29% chi squared 4.24 p=.04).

The present: Prospective Outcome Study

In 1990, a large prospective outcome study was launched by the then single-handed researcher, Dr. Bridget Dolan, who was at the time based solely in the hospital with no other formal academic links. Given the slender resources, this was an ambitious project. The study attempted to produce a psychological profile of all referrals during the study period. In addition it aimed to rate all

patients admitted at 3 monthly intervals and at one and three years post-discharge; and those not admitted at one and three years post-referral. At this time there was no published large-scale prospective study in the field of personality disorder and the proposed study was breaking new ground in terms of improved research methodology.

Shortly after the study began, in 1991, major changes to the UK's National Health Service took place. In effect, service provision of the NHS was separated organisationally from its purchasing. As a consequence of the changes, local districts became more responsible for identifying the health needs of 'their' catchment area populations and using their funding allocation from central Government to purchase these. Henderson's national catchment area meant that its funding depended on referrals other than from its local catchment area ('extra-contractual referrals'). Henderson's client group did not compete well with other extra-contractual referrals in the new 'market-place' (Dolan et al, 1994).

The numbers of applications for admission that attracted funding reduced, the financial viability of the hospital was threatened as, in the minds of purchasers, the treatment was not considered of proven worth but an expensive luxury which NHS could no longer afford. Paradoxically, two research benefits emerged from this otherwise negative scenario. First, a study of cost-offset was evoked in order to challenge the 'expensive' price tag (Menzies et al. 1993); (Dolan et al. 1996). Second, a group of referrals emerged who did not get admitted for treatment solely because the health authority refused to fund their ECR. This, it could be argued provided a comparison group closer to a randomly allocated control than the whole group of non-admitted referrals, for whom non-admission may have resulted from characteristics also relevant to treatment outcome (Dolan et al. 1997).

Cost as a measure of outcome

Two papers report the cost-offset of Henderson treatment. In the first study service usage data were collected retrospectively from May 1992 on 29 consecutive admissions and success rates of previously published outcome studies were used to calculate cost offset. In the second study actual service usage in the one-year post treatment was used to calculate the actual cost-offset for these admissions.

Data on mental health and forensic service usage in the one year prior to admission to Henderson derived from three sources: (1) case notes including information provided by the referrer; (2) the "Social History Form", a questionnaire completed by all admissions, concerning family, personal and

clinical history; (3) subjects who were resident during the study period completed an additional form about the previous year's usage of services. Costs were calculated by obtaining figures of extra contractual referrals (ECR) tariffs for 1992/3 from the then four Thames Regional Health Authorities (RHA).

The daily tariff for Henderson Hospital was £111 compared with £153.20 for a general acute psychiatric bed and £173 for Close Supervision Units. The twenty-nine Henderson admissions had used a considerable amount of health and prison services in the previous year, the average estimated costs were £423,115 per year (mean cost per patient £14,590). In this first study, the cost of treatment at Henderson was offset by extrapolating from the 41% success rate in the studies reviewed above. This suggested that the treatment would pay for itself in four years (Menzies et al, 1993).

Subsequently in the second study, follow-up information had been obtained in the course of the prospective outcome study on actual service usage for 24 (83%) of the 29 residents in the original sample (Menzies et al, 1993). The average cost of services used by these twenty-four residents in the one year prior to admission was £13,966.

Information on service usage in the one year subsequent to discharge from Henderson was obtained from their referrer (in 17 cases) and/or their G.P. (in 14 cases) and from the client themselves (in 7 cases). Four subjects had further in-patient admissions, one of whom was re-admitted to the Henderson. Two people had outpatient assessments, twelve had outpatient treatment and one attended a day hospital. None of the residents spent time in prison or a secure psychiatric unit during the year. The average cost of services used was £1,308. This represents an annual saving post-discharge of £12,658.

These 24 residents were in treatment at Henderson Hospital for an average of 231 days (range 1 to 365 days) thus the actual cost of their treatment at Henderson was £25,641. Should the cost-offset continue at a similar rate for subsequent years then the cost of admission to Henderson would be re-couped in less than two years and represent savings thereafter.

Borderline symptoms as a measure of outcome

Reviewing outcome studies in the field of personality disorder (Norton and Dolan 1995c) note that many studies fail to assess the impact of treatment on aspects intrinsic to the personality disorder pathology itself, separately from those that are only associated or indirect phenomena. Indeed, there is a range of features associated with personality disorder, changes in which are erroneously equated with change in the personality disorder itself, such

as reduction in axis 1 diagnosis symptomatology, or behavioural features such as criminal activity, self-mutilation or suicidality. This prospective study aimed to assess changes in core personality disorder features. Comparison was made between those admitted and those not admitted for treatment. Consecutive referrals were mailed a self-report questionnaire pack on referral, including the Personality Disorder Questionnaire (PDQ-R) (Hyler et al. 1987), the Borderline Syndrome Index (BSI) (Conte et al. 1980), the Irritability, Depression and Anxiety Scale (IDA) (Snaith et al. 1978), and the Rosenberg Self-Esteem Scale (RSE) (Rosenberg 1965). A second follow-up assessment pack was set one year after referral (for the not-admitted) or discharge (for the admitted group). Up to three repeated mailings were used to maximise response rate.

The results of this study showed a significantly greater reduction in BSI scores in the treated than in the non-admitted group (Dolan, 1997). Changes in BSI scores were significantly positively correlated with length of treatment in the admitted group. Again, assessment of the reliability and clinical significance (Jacobson and Truax 1991) of changes in individual subjects was conducted in this study. These showed that the magnitude of this change was reliable and clinically significant in 42.9% of the admitted sample, compared with only 17.9% of the non-admitted sample (18.2% of the unfunded group).

Mood as a measure of outcome: Work in progress

Further results from the prospective outcome study described above are reported here. Given the positive effects shown on Borderline symptomatology, the outcome in terms of mood was also of interest. Depression, anxiety and inwardly and outwardly directed irritability were assessed using the IDA (Snaith et al. 1978) on the same cohort of patients described above. This instrument was chosen because it covers relevant symptoms theoretically related to personality disorder: anxiety and depression and, in addition, assesses the socially relevant tendencies to hurt the self or others. The scale is an 18 item scale with items scored from 0-4. The depression and anxiety subscales are scored between 0 and 15 and the irritability scales are scored between 0 and 12.

Participants

Response rates

The study sample is derived from the sample on which borderline symptoms were reported by Dolan, Warren et al. (1997). However the study period was extended therefore the sample and response rates will be summarised here. Consecutive referrals to the service between September 1990 and December 1994 were approached to participate in the study. There were 585 eligible referrals in the study period. Three hundred and eighty-four referrals (66%) completed a baseline assessment. Twelve of these participants (3%) were excluded from follow-up because they were rereferred for treatment in the period between

initial and follow-up assessment. One hundred and thirty five (36%) of the 372 eligible, responded at follow-up assessment. Seventy-five (56%) of the follow-up respondents were admitted and 60 (44%) not admitted.

Demographic characteristics

The mean age of the sample was 28 (range, 17-49, sd=6.8). Just under one half was female. Almost all the sample was white, single and unemployed. A large proportion had been previously convicted and had histories of drug and alcohol abuse. There were also high rates of previous suicide attempts, self-mutilation and overdosing reported by the referrers.

Personality disorder

This sample is a severely personality disordered group of people. The mean number of personality disorders for which each individual met criteria was seven and 95% of the sample met criteria for more than one personality disorder. Two cases did not score above threshold for any personality disorder and one case had missing data. The mean PDQ-R total score was 58 (see Table 1).

Table 1 Personality disorder symptoms: breadth and number of diagnoses

	Participants (n=134)
PDQ-R total score	
Mean (range), sd	**57.49 (17-86), 13.00**
Number of diagnoses[11]	
Mean (range), sd	**6.80 (0-12), 2.94**
More than one diagnosis	
Number of participants (%)	**127 (94.8)**

Eighty-four per cent of these participants met criteria for Borderline Personality Disorder.

Table 2 shows the prevalence of personality disorder diagnoses. Following Borderline, the most common diagnoses are Paranoid and Schizotypal, and Avoidant. The two research categories (self-defeating and sadistic) are the least prevalent.

Table 2 Personality disorder sub-category diagnoses

[11] The research categories are included so that this is the number of diagnoses out of a possible 13

Sub-category diagnoses	% Admitted (n=74)	% Not-admitted (n=60)
Paranoid	**77**	**65**
Schizoid	**55**	**33***
Schizotypal	**74**	**62**
Antisocial	**58**	**67**
Borderline	**82**	**85**
Histrionic	**55**	**67**
Narcissistic	**39**	**40**
Avoidant	**69**	**62**
Obsessive	**47**	**38**
Dependent	**62**	**52**
Passive-aggressive	**47**	**33**
Self-defeating	**14**	**22**
Sadistic	**22**	**27**

*** significant at p<.05**

The most prevalent diagnoses were spread across the putative clusters into which personality disorder diagnoses are grouped in DSM-III-R & DSM-IV. Each cluster contained at least one score for over 80% of the group. Over 70% of the participants met criteria for personality disorders in all three clusters. There was a slightly higher proportion of Schizoid personality disorder in the admitted than the non-admitted group.

Representativeness

Given the naturalistic nature of this study and the attrition of participants over time, tests were conducted to establish the representativeness of the sample. Group differences at baseline suggested that outcome analyses should take sex and the presence of schizoid personality disorder.

Length of stay

The average length of stay of the admitted participants in this comparison was 201 days (6.7 months). The minimum length of stay was 2 days and the maximum 396.

Outcomes

The results of repeated measures analysis of variance are summarised in Table 3. The table shows the baseline and follow-up mean score for each group and the interaction effect. The BSI results have been included here to provide continuity with the previously published results on this measure (Dolan et al. 1997). A highly significant interaction effect is shown.

Table 3 Summary of Outcomes

		Baseline	*One year*	F	P
Borderline Symptoms	Admitted	34.89 (9.39)	22.03 (15.29)	10.85	.001
	Not admitted	32.98 (11.55)	28.26 (12.90)		
Anxiety	Admitted	10.01 (2.73)	7.61 (4.04)	11.98	.001
	Not admitted	8.67 (3.07)	8.56 (3.44)		
Depression	Admitted	8.73 (3.09)	6.70 (3.98)	4.61	.034
	Not admitted	7.68 (3.78)	7.14 (3.13)		
Inward Irritability	Admitted	8.48 (3.28)	6.25 (3.73)	7.78	.006
	Not admitted	6.80 (3.53)	7.14 (3.13)		
Outward Irritability	Admitted	6.53 (2.78)	5.77 (3.16)	5.09	.026
	Not admitted	6.33 (2.87)	6.76 (3.02)		

The results show highly significant interactions for anxiety, depression, inwardly and outwardly directed irritability. The mean scores show that in each case this interaction is a result of greater improvement in the admitted group between baseline and one year follow-up. There is a suggestion of improvement in the non-admitted group for Borderline symptoms, anxiety and depression, but of a deterioration in inward and outwardly directed irritability.

The group differences at baseline on anxiety and inward irritability were significant with higher (more pathological) scores in the admitted group. Adjusting for this produced a non-significant interaction for inwardly directed irritability. There was a main effect of time, significant at the $p<.05$ level, however, suggesting that both groups were showing some improvement on this measure.

Confounding variables

When sex, and Schizoid Personality Disorder were entered as factors into the individual analyses of variance, no interactions or main effects of these variables were found. For outwardly directed irritability, there was no interaction of schizoid personality disorder with either time or group.

Relationship of outcomes to length of stay in treatment - Association of length of stay with follow-up score

Length of stay was negatively correlated with all follow-up scores: the longer a resident stayed in treatment the lower their follow-up score. These negative correlations were significant to $p<.05$ level with follow-up scores on borderline symptoms, anxiety, depression. The negative correlations with inwardly and outwardly directed irritability were not significant.

Association of length of stay with change

(Dolan et al. 1997) found a significant correlation between change in borderline symptoms and length of stay in treatment. These analyses also show a significant correlation between length of stay in treatment and degree of change between baseline and one-year post treatment follow-up for depression and anxiety. However, there was no signifcant correlation for inwardly and outwardly directed irritability.

Comparison of change for short-stay and long-stay participants

The admitted patients were therefore divided into long and short stay groups. Short-stayers were those who stayed less than three months and long-stayers those who stayed nine months or more in treatment. The short stay group stayed in treatment a mean 34 days (sd 29.9, range 2-91 days). The long stay group remained in treatment a mean of 343 days (sd 32.2, range 277-396 days).

T-tests revealed significant differences for the short and long stay groups in change scores for borderline symptoms, anxiety, depression, and inwardly directed irritability. All of changes are improvements for the admitted group. Differences in change scores were non-significant for outwardly directed irritability.

Earlier studies (Dolan, 1992; Copas et al. 1984) had found a significant gender difference in length of stay. This was not evident in this study.

Summary of work in progress

These results augment the existing evidence of positive treatment outcomes. Improvements in core personality disorder pathology previously shown (Dolan et

al. 1997) are supported by the improvements in mood symptomatology shown here. Treatment effects seem to be shown in terms of reductions in anxiety, depression and outwardly directed irritability, although reduction in inwardly directed irritability would seem to be a weaker effect. It is of interest that there were significant differences between the admitted and non-admitted referrals in terms of anxiety and inwardly directed irritability on which the admitted group scored more highly. This may suggest a selection effect, which should be explored by future research.

Comment

This prospective outcome study also suffers from some of the methodological shortcomings levelled at previous studies, above. The use of self-report measures limits the validity of the findings, although some reassurance can be gained from the consistency of effect across multiple self-report measures. The proportion of those about whom we have data at outcome is only a small proportion of the eligible sample in both groups. However, the response rate in this study is not atypical of a PD sample of patients. This limits the generalisability of the findings to PD patients in general. In addition, the follow-up interval differed between the treated and non-treated samples so that the non-treated sample was followed up earlier than those treated. Alternative study designs, which attempted to match the timing of a non-treated follow-up with a treated follow-up, would only have been possible in theory, since the time spent in treatment for any individual patient could not be known in advance! However, it is also highly unlikely that time alone accounts for such a magnitude of difference between the admitted and non-admitted groups when spontaneous remission in this client group is widely acknowledged to be rare. Further analysis of results using the data collected during treatment may help to substantiate this. The non-admitted comparison group cannot truly be labelled "non-treated" since it is likely that they had at least some non-specialist treatment during the study period which could not be controlled for. The use of this comparison sample is also problematic because the reasons for non-selection or non-attendance may relate to a poorer prognosis at the point of entry to the study. The use of the non-funded group in the study on borderline symptoms, however, provides some reassurance. The non-random allocation to treatment or non-treatment, is the most problematic methodological limitation of the study because this allows systematic variation between groups. Some of the difficulties of applying randomisation in this context are identified in the discussion below.

Summary of outcome studies of Henderson Hospital

Table 4 shows a summary of the outcome studies reviewed above. These have shown improvements in those admitted for treatment using a range of approaches to outcome measurement including convictions, psychiatric service usage and a various kinds of psychological functioning. The proportion improved seems to be consistently around 40% but this may improve to around 70% for those who stay in treatment for nine months or more. Treatment gains have been shown to

be maintained up to five years post treatment. There is some evidence that a small proportion 3-35% deteriorate following treatment. Where comparison groups have been used, they have been shown to fare significantly worse than those admitted to the treatment although a small proportion of “untreated” controls also show improvement over time. Some of the earlier studies were limited in terms of comparison groups, measures used and follow-up periods although the methodological approaches can be seen to have evolved over time.

Table 4 Summary of outcome studies of Henderson Hospital treatment

STUDY	N	FOLLOW-UP PERIOD	CRITERIA OF SUCCESS	
Rapoport (1960)	64	1 year	Improved clinically since admission	
Tuxford (1961)	86	2 years	In employment No recidivism	
Taylor (1963)	?	9 months	In employment	
Whiteley (1970)	112	2 years	No recidivism No re-admission Neither of above	
Copas and Whiteley (1976)	104 87 104	2 years 5 years	No recidivism or re-admission “	
Copas et al. (1986)	194 51	3 years 5 years 3 years 5 years	“ “ “ “	
Dolan et al. (1991)	62	8 months	Improved psychological functioning on SCL-90	
Dolan, Warren, Menzies & Norton (1996)	24	1 year	Group cost reduction	

Dolan, Warren & Norton (1997)	70 67	1 year	Clinically significant change in Borderline Personality Disorder symptomatology than in untreated controls	

Table 5 Summary of outcome studies of Henderson Hospital treatment

General issues for research in this context:

Randomised controlled trials

None of these outcome studies utilised a random allocation to treatment or control group. Randomisation is a problematic option for many reasons in this context. In the early life of the Unit not only were RCTs rare, even in the field of psychopharmacology but also in other simpler experimental situations in which a clear outcome result could be swiftly obtained, in a matter of days or weeks. However, what is acceptable methodologically has changed since the 1940s and 50s and more recently, the advent of initiatives such as Evidence Based Medicine (EBM) has increased the profile of RCT evidence for treatments. It is important to discuss the issues involved in applying RCT methods to therapeutic community treatment for this reason.

Residential therapeutic community treatment represents both a complex and lengthy treatment (average duration 7 months) and personality disorder is a complex problem, for which there is no agreed single outcome measure but for which several domains, social, psychological and penal are relevant. The stay of a given patient in the therapeutic community may be up to one year and the nature of the treatment requires a community of up to a maximum of 29, and sustained change post-discharge is needed to provide evidence of an actual treatment effect. Although all RCTs are resource intensive, an RCT of this treatment would be particularly so and in previous years funding has not been available to support such a study.

However, there are more difficult issues in this context than lack of finances. The democratic TC wherein patients treat one another through an exclusively group treatment approach requires that an adequate number of community members is maintained. From both a clinical and research perspective this will always remain a central concern. Since any imposed research constraint, such as randomisation could threaten the integrity of the treatment under study if it

resulted in inadequate numbers of patients being referred to, or agreeing to be allocated randomly to, the treated sample at a particular point in time.

Some of the strength of this argument against randomisation, would be lessened were there to be a sufficient and guaranteed excess of referrals over treatment places such as if the waiting list for admission were sufficiently long. However, this raises two issues. First, a key feature of this treatment is the process of selecting new residents. The TC requires that patients are involved in the selection of new residents. In fact, the resident group has the majority vote in decision-making about selection candidates. This requirement increases the resident community's responsibility for the decision to admit an individual. For example, this reduces the likelihood of the resident community "washing its hands" of a new resident who turns out to be unpopular and handing over the problem to "staff" who are seen to have imposed this person on the community. In addition, both the community and the new resident have some knowledge of each other before admission, a process which may increase engagement through the candidate's identification with other similar people already receiving treatment in this community. Where there is a waiting list of months duration this requirement is decreasingly met since those who selected the patient are increasingly likely to have themselves left the TC, hence the authenticity and power of the treatment is likely to be diminished. Secondly, this patient group is at high risk of self-damaging behaviour, reconviction of offending or changes of heart or circumstances which suggest opportunities to engage them in treatment should be taken as quickly as possible when they arise. An ethical dilemma is presented by augmenting the waiting list and hence the wait for treatment.

Another difficulty for randomisation at the beginning of the referral process arises from selection. Whilst the highly selected nature of the patient group has been levelled as a criticism of the naturalistic research conducted thus far, it is viewed, from a clinical point of view as an essential part of the treatment. The specialist nature of the treatment entails that it is not suitable for all who suffer from the condition and a schematic process to avoid admitting those who might be damaged is therefore required. For these reasons randomisation would need to occur post-selection. This is likely to have major negative implications for the enhancement of engagement promoted by the selection process as described above.

In the absence of alternative proven effective treatment for people with difficulties this severe there is a clear ethical obstacle. Those who refused to

consent to the randomised study conditions to be refused treatment. If there was a ‘mixture’ of randomised and non-randomised this could pose problems, certainly for the maintenance of ‘blindness’. These are just some of the issues pertinent to RCTs of this treatment. These would be in addition to the other problems with implementing RCTs on long-term treatments for personality disorder with low throughput. As yet, research has not overcome the difficulties posed by the RCT design.

Future directions

From the results of these outcome studies there is a suggestion that the intervention of democratic therapeutic community treatment might be of significant benefit to at least some patients with severe personality disorder. A positive association between length of treatment received and improved outcome has been repeatedly described although this has not always been statistically significant. When “non-treated” samples have been compared, there has been a consistent finding of significantly greater improvements following this specialist treatment. Future research should broaden the domains in which outcome is measured and look for overall treatment effects. Further consideration should be given to the applicability of the RCT design and the selection process should be explored further in order to identify more clearly those patients who can be expected to benefit from this treatment and those who cannot.

References

Conte, H. R., Plutchik, R., Karasu, T. B., et al. A self-report borderline scale: discriminant validity and preliminary norms. *Journal of Nervous and mental disease* (1980) 168(7):428-435.

Copas, J., and Whiteley, J. S. Predicting success in the treatment of psychopaths. *British Journal of Psychiatry* (1976) 129:388-92.

Copas, J. B., O'Brien, M., Roberts, J. C., et al. Treatment outcome in personality disorder: The effect of social, psychological and behavioural variables. *Personality and Individual differences* (1984) 5:565-73.

Derogatis, L. R., Rickels, K., and Rock, A. F. The SCL-90 and the MMPI: A Step in the Validation of a New Self-Report Scale. *British Journal of Psychiat* (1976) 128:280-289.

Dolan, B., and Coid, J. *Psychopathic and Antisocial Personality Disorders: Treatment and Research Issues*. Gaskell, 1994.

Dolan, B., Morton, A., and Wilson, J. Selection of admissions to a therapeutic community using a group setting: Association with degree and type of psychological distress. *International Journal of Social Psychiatry* (1990) 36:265-271.

Dolan, B., Warren, F., Menzies, D., et al. Cost-offset following specialist treatment of severe personality disorders. *Psychiatric Bulletin* (1996) 20:413-17.

Dolan, B., Warren, F., and Norton, K. Change in Borderline Symptoms One Year After Therapeutic Community Treatment For Severe Personality Disorder. *British Journal of Psychiatry* (1997) 171:274-279.

Dolan, B. M., Evans, C., and Wilson, J. Therapeutic community treatment for personality disordered adults: Changes in neurotic symptomatology on follow-up. *International Journal of Social Psychiatry* (1992) 38:243-250.

Dolan, B. M., Evans, C. D., and Norton, K. Multiple Axis-II diagnosis of personality disorder. *British Journal of Psychiatry* (1995) 106:107-112.
Henderson, D. K. *Psychopathic states*. Chapman & Hall, 1939.
Hyler, S., Reider, R. O., Spitzer, R. L., et al. *Personality Diagnostic Questionnaire - revised (PDQ-R)*. New York State Psychiatric Institute, 1987.

Jacobson, N. S., and Truax, P. Clinical Significance: a Statistical Approach to Defining Meaningful Change in Psychotherapy Research. *Journal of consulting and clinical psychology* (1991) 59:12-19.

Jones, M. *A study of therapeutic communities*. Tavistock, 1952.

Main, T. F. The hospital as a therapeutic institution. *Bulletin Menninger Clinical* (1946) 10:66-68.

Main, T. F. The Ailment. *British Journal of Medical Psychology* (1957) 30:129-145.

Manning, N., and Rapoport, R. Rejection and Reincorporation: A case study in social research utilisation. *Social Science and Medicine* (1976) 10:459-468.

Menzies, D., Dolan, B. M., and Norton, K. Are short term savings worth long term costs? Funding psychotherapeutic inpatient treatment for personality disorders. *Psychiatric Bulletin* (1993) 17:517-519.

Norton, K. A culture of enquiry - its preservation or loss. *International Journal of Therapeutic Communities* (1992a) 13(1):3-26.

Norton, K. How Therapeutic Communities Work. In G. Edwards and C. Dare, eds., *Psychotherapy Psychological Treatment and the Addictions*. Cambridge University Press, 1996.

Norton, K., and Dolan, B. Acting out and the institutional response. *Journal of Forensic Psychiatry* (1995a) 6(2):317-332.

Norton, K., and Dolan, B. Assessing change in personality disorder. *Current Opinions in Psychiatry* (1995b) 8:371-375.

Norton, K., and Hinshelwood, R. D. Severe Personality Disorder: Treatment Issues and Selection for In-Patient Psychotherapy. *British Journal of Psychiatry* (1996) 168:723-731.

Norton, K. R. W. Personality disordered individuals: the Henderson Hospital model of treatment. *Criminal Behaviour and Mental Health* (1992b) 2:180-191.

Norton, K. R. W. Personality disordered forensic patients and the therapeutic community. In C. Cordess and M. Cox, eds., *Forensic Psychotherapy*. Jessica Kingsley Publishers, 1995.

Norton, K. R. W., and Dolan, B. Assessing Change in Personality Disorder. *Current Opinions in Psychiatry* (1995c) 8:371-375.

Rapoport, R. *The community as doctor*. Tavistock, 1960.

Rosenberg, M. The Measurement of Self-Esteem. *Princeton Univ Press Princetown Princetown* (1965).

Rosenbluth, M. New uses of countertransference for the inpatient treatment of borderline personality disorder. *Canadian Journal of Psychiatry* (1991) 36:280-4.

Snaith, R. P., Constantopoulos, A. A., Jardine, M. Y., et al. A clinical scale for the self assessment of irritability. *British Journal of Psychiatry* (1978) 132:164-171.

Taylor, F. H. The Treatment of Delinquent Psychopaths. *The Howard Journal* (1963) 1:1-9.

Tuxford, J. Treatment as a circular process. *King Edward's Hospital Fund London* (1961).

Tyrer, P. J. T., and Johnson, T. Establishing the Severity of Personality Disorder. *American Journal of Psychiatry* (1996) 153(12):1593-7.

Wesby, R., Menzies, D., Dolan, B., et al. A survey of psychological types in a therapeutic community. *Therapeutic Communities* (1995) 16:229-238.

Whiteley, J. S. The Henderson Hospital: A community study. *International Journal of Therapeutic Communities* (1980) 1(2):38-58.

Whiteley, J. S. Sociotherapy and Psychotherapy in the treatment of personality disorder. Discussion paper. *Journal of the Royal Society of Medicine* (1986) 79:721-5.