Maarit Laaksonen

Patient suitability for short-term and long-term psychotherapy

144

Maarit Laaksonen

Patient suitability for short-term and long-term psychotherapy

ACADEMIC DISSERTATION

To be publicly discussed, by the permission of the Faculty of Social Sciences of the University of Jyväskylä, in Mattilanniemi, Auditorium MaA 211, on December 17, 2014 at 12 o'clock noon.

National Institute for Health and Welfare, Helsinki, Finland
and
Department of Psychology, University of Jyväskylä, Finland

Helsinki 2014

NATIONAL INSTITUTE
FOR HEALTH AND WELFARE

Cover photo: Timo Heino: Scar (Detail of the installation), 2010. Kerava Art Museum, Pain exhibition. Photograph: Petri Virtanen.

Supervisors
Professor Emeritus Paul Knekt, PHD
Department of Health, Functional Capacity and Welfare
National Institute for Health and Welfare
Helsinki, Finland

Dr Olavi Lindfors, PhD
Department of Health, Functional Capacity and Welfare
National Institute for Health and Welfare
Helsinki, Finland

Professor Jarl Wahlström, PhD
Department of Psychology
University of Jyväskylä
Jyväskylä, Finland

Reviewers
Professor Emeritus Juhani Julkunen, PhD
Institute of Behavioural Sciences
University of Helsinki
Helsinki, Finland

Docent Risto Antikainen, PhD
Department of Psychology
University of Jyväskylä
Jyväskylä, Finland

Opponent
Professor Emeritus Johannes Lehtonen, PhD
University of Eastern Finland
Finland

"We'll do it all
  Everything
  On our own

  We don't need
  Anything
  Or anyone

  If I lay here
  If I just lay here
  Would you lie with me
  And just forget the world?"

From "Chasing Cars" by Snow Patrol

To my Mum

Abstract

A clinically-informed, evidence-based choice of optimal psychotherapy is considered of the utmost importance to patients' recovery from mental disorders. Nevertheless, no generally accepted, systematic method of referring psychiatric patients to psychotherapy currently exists. Patients' psychological, intrapsychic and interpersonal, capacities are among the core psychotherapy suitability selection criteria suggested for differentiating those treatable with short-term psychotherapy from those in need of long-term psychotherapy. Research into both the assessment of psychotherapy suitability and its prediction of psychotherapy outcomes remains scarce, however. Only a few psychotherapy suitability assessment scales have been studied for their reliability and validity. Psychotherapy outcome prediction by patient suitability has been almost exclusively studied based on short-term therapies. No study has compared outcome prediction by patient suitability in short-term versus long-term psychotherapy. In this study, a new interview-based 7-item Suitability for Psychotherapy Scale (SPS) was presented and its reliability, validity, and prediction of outcomes of short-term and long-term psychotherapy were thoroughly evaluated based on data from the Helsinki Psychotherapy Study (HPS).

In the HPS, 326 outpatients aged 20-46 years and suffering from mood and anxiety disorders were randomized between individual short-term, psychodynamic (N = 101) or solution-focused (N = 97), psychotherapy or long-term psychodynamic psychotherapy (N = 128). Prior to randomization, the patients' psychological capacities (i.e., modulation of affects, flexibility of interaction with the interviewer, self-concept in relation to ego ideal, reflective ability, response to trial interpretation, motivation, and the focality of the problems) were assessed using the SPS. A cumulative SPS score, varying between 0-7, was formed by summing up the values of the seven suitability measures, classified as good (0) or poor (1). Psychiatric symptoms were self-reported using the Symptom Checklist-90, Global Severity Index (SCL-90-GSI) at baseline and seven times during a 3-year follow-up. The reliability of the SPS was evaluated based on a non-random sample of 39 videotaped interviews by measuring the agreement between the seven individual interviewers' assessments and the repeatability of the interviewers' assessments over three years using kappa coefficients. The validity of the SPS was evaluated based on the total sample of 326 patients by measuring its association with a criterion measure, Quality of Object Relations Scale (QORS), and a discriminating measure, SCL-90-GSI, at baseline using linear models. Prediction by the seven SPS measures and the SPS score of the development of psychiatric symptoms in the two short-term

therapies and the long-term therapy during the 3-year follow-up was measured using linear mixed models, adjusting for baseline symptoms.

The findings supported both the agreement between individual interviewers' assessments and the repeatability of the interviewers' assessments over time. An association of the SPS with the personality functions but not with psychiatric symptoms supported the criterion and discriminating validity of the SPS. The SPS predicted changes in psychiatric symptoms during follow-up for all three therapy groups irrespective of the baseline symptom level. The SPS did not notably differentiate the outcome of the two short-term therapies, but strongly differentiated the outcome of short-term and long-term psychotherapy. Patients with good values in the individual SPS measures experienced a faster symptom reduction in short-term therapy, whereas patients with poor values in these measures experienced a greater symptom reduction in long-term therapy. Three patient groups with different prognosis were identified when the cumulative SPS score was used to predict symptom development: 1) patients with mainly good values in the seven suitability measures (score values 0-3) seemed to benefit more from short-term therapy, 2) patients with mainly poor values in the seven suitability measures (score values 4-6) seemed to benefit more from long-term therapy, and 3) patients for whom all seven values were poor (score value 7) seemed to fail to benefit from either short-term or long-term therapy.

The SPS appeared to be a reliable and valid pre-treatment psychotherapy suitability assessment method, with the ability to predict and differentiate the outcome of short-term and long-term psychotherapy. The SPS may thus be utilized in matching patients to treatments and optimizing both treatment outcomes and health care resources. However, more research is needed to confirm these findings and to demonstrate their usefulness in practice.

Tiivistelmä

Maarit Laaksonen. Patient suitability for short-term and long-term psychotherapy. [Potilaiden soveltuvuus lyhyeen ja pitkään psykoterapiaan]. Terveyden ja hyvinvoinnin laitos. Tutkimus 144, 206 sivua. Helsinki, Finland 2014.
ISBN 978-952-302-332-1 (painettu); ISBN 978-952-302-333-8 (pdf)

Kliiniseen kokemukseen ja tutkimusnäyttöön perustuva optimaalisen psykoterapiamuodon valinta on tärkeä edellytys mielenterveyshäiriöistä parantumiselle. Mitään yleisesti hyväksyttyä, yhtenäistä psykiatristen potilaiden hoitoonohjauskäytäntöä ei ole kuitenkaan olemassa. Potilaan persoonallisuuteen ja interpersoonallisiin taipumuksiin liittyvää psykologista kyvykkyyttä on pidetty olennaisena arvioitaessa ja eroteltaessa lyhyen ja pitkän psykoterapian soveltuvuutta ja tuloksellisuutta. Sekä psykologisten potilastekijöiden luotettavaan arviointiin että ennustekykyyn liittyvä tutkimus on kuitenkin vähäistä. Psykoterapiasoveltuvuuden arviointiasteikkojen reliabiliteettia ja validiteettia on tutkittu vain harvoin, ja niiden ennustekykyä on arvioitu lähinnä suhteessa erityyppisten lyhyterapioiden tuloksellisuuteen. Psykologisten potilastekijöiden ennustekykyä lyhyessä ja pitkässä yksilöpsykoterapiassa ei ole toistaiseksi verrattu. Tässä tutkimuksessa arvioitiin uuden haastatteluun pohjautuvan 7-osioisen Psykoterapiasoveltuvuuden arviointiasteikon (Suitability for Psychotherapy Scale, SPS) reliabiliteettia, validiteettia ja ennustekykyä suhteessa lyhyen ja pitkän psykoterapian tuloksellisuuteen, Helsingin Psykoterapiatutkimuksen aineistoon pohjautuen.

Helsingin Psykoterapiatutkimuksessa satunnaistettiin 326 iältään 20-46 –vuotiasta, masennuksesta tai ahdistuneisuushäiriöstä kärsivää avohoitopotilasta joko lyhyeen, psykodynaamiseen (N = 101) tai voimavarasuuntautuneeseen (N = 97), tai pitkään, psykodynaamiseen (N = 128), terapiaan. Alkuhaastattelussa, ennen potilaiden satunnaistamista eri terapiamuotoihin, arvioitiin SPS-asteikon avulla heidän psykologista kyvykkyyttään: tunteiden hallintaa, vuorovaikutuksen sujuvuutta, minäkäsityksen suhdetta ideaaliminään, reaktiota koetulkintaan, reflektointikykyä, motivaatiota ja ongelmien fokusta. Näiden seitsemän soveltuvuusmuuttujan arvot luokiteltiin hyviksi (0) tai huonoiksi (1) ja ne summaamalla muodostettiin kullekin potilaalle kokonaissoveltuvuusindeksi (0-7). Potilaiden psykiatrisia oireita arvioitiin SCL-90 –oirekyselyllä alkumittauksessa ja seitsemän kertaa kolmen vuoden seurannan aikana. SPS-asteikon reliabiliteettia arvioitiin mittaamalla seitsemän haastattelijan 39 potilaan videoitujen haastattelujen perusteella tekemien soveltuvuusarviointien yhtäpitävyys ja toistettavuus kappakertoimen avulla. SPS-asteikon kriteerivaliditeettia arvioitiin mittaamalla sen yhteyttä toiseen psykologista kyvykkyyttä, objektisuhteiden laadun näkökulmasta, mittaavaan QORS (Quality of Objects Relations Scale) –asteikkoon ja erotteluvaliditeettia mittaamalla sen yhteyttä SCL-90 –oirekyselyyn, lineaaristen mallien avulla, 326 potilaan alkumittauksiin perustuen. SPS-asteikon seitsemän

yksittäisen soveltuvuusmuuttujan ja kokonaissoveltuvuusindeksin ennustekykyä suhteessa potilaiden psykiatristen oireiden kehittymiseen kahdessa lyhyessä ja yhdessä pitkässä terapiassa kolmen vuoden seurannan aikana mitattiin lineaaristen sekamallien avulla, potilaiden alkuoiretaso vakioiden.

Tämän tutkimuksen tulokset tukivat sekä haastattelijoiden tekemien SPS-soveltuvuusarviointien yhtäpitävyyttä että niiden toistettavuutta kolmen vuoden välein. SPS-asteikon ja QORS-asteikon välinen voimakas yhteys tuki kriteerivaliditeettia ja SPS-asteikon ja SCL-90 –oirekyselyn välinen heikko yhteys erotteluvaliditeettia. SPS-asteikko ennusti potilaiden seurannan aikaisten oireiden kehittymistä, heidän alkuoiretasostaan riippumatta, kaikissa kolmessa tutkitussa terapiamuodossa. SPS-asteikko ei juurikaan erotellut kahden eri lyhytterapian välistä tuloksellisuutta, mutta erotteli voimakkaasti lyhyen ja pitkän terapian välistä tuloksellisuutta. Potilaat, joiden soveltuvuustekijät oli arvioitu hyviksi, kokivat oireiden vähenevän nopeammin lyhyessä kuin pitkässä terapiassa, kun taas potilaat, joiden soveltuvuustekijät oli arvioitu huonoiksi, kokivat pitkällä tähtäimellä oireiden vähenevän enemmän pitkässä terapiassa. Potilaiden oireiden kehitystä kokonaissoveltuvuusindeksillä ennustettaessa löydettiin kolme erilaista potilasryhmää: 1) potilaat, joilla enemmistö seitsemästä soveltuvuustekijästä oli arvioitu hyviksi (soveltuvuusindeksin arvot 0-3), näyttivät hyötyvän enemmän lyhyestä kuin pitkästä terapiasta, 2) potilaat, joilla enemmistö seitsemästä soveltuvuustekijästä oli arvioitu huonoiksi (soveltuvuusindeksin arvot 4-6), näyttivät hyötyvän enemmän pitkästä kuin lyhyestä terapiasta ja 3) potilaat, joilla kaikki seitsemän soveltuvuustekijää oli arvioitu huonoiksi, eivät näyttäneet hyötyvän lyhyestä eivätkä pitkästä terapiasta.

SPS-asteikko osoittautui tässä tutkimuksessa luotettavaksi ja päteväksi psykoterapiasoveltuvuuden arviointimenetelmäksi, jolla voitiin ennustaa ja erotella lyhyen ja pitkän terapian tuloksellisuutta. Tätä asteikkoa voitaneen siis hyödyntää potilaalle parhaiten soveltuvan psykoterapian valinnassa ja siten hoitotulosten ja hoitoresurssien optimoinnissa. Jatkotutkimuksia kuitenkin tarvitaan näiden tulosten vahvistamiseksi ja SPS-asteikon käytännön hyödyllisyyden osoittamiseksi.

Avainsanat: ahdistuneisuushäiriö, ennuste, lyhyt psykoterapia, mielialahäiriö, pitkä psykoterapia, psykologinen kyvykkyys, psykoterapiasoveltuvuuden arviointi, reliabiliteetti, validiteetti

Contents

List of original papers

I   Laaksonen MA, Lindfors O, Knekt P, Aalberg V. Suitability for Psychotherapy Scale (SPS) and its reliability, validity, and prediction. *British Journal of Clinical Psychology* 2012; 51: 351–375.

II   Laaksonen MA, Knekt P, Sares-Jäske L, Lindfors O. Psychological predictors on the outcome of short-term psychodynamic psychotherapy and solution-focused therapy in the treatment of mood and anxiety disorder. *European Psychiatry* 2013; 28: 117-124.

III   Laaksonen MA, Knekt P, Lindfors O. Psychological predictors of the recovery from mood or anxiety disorder in short-term and long-term psychotherapy during a 3-year follow-up. *Psychiatry Research* 2013; 208: 162-173.

These articles are reprinted with the kind permission of their copyright holders.

Abbreviations

ACT	Acceptance and Commitment Therapy
ASQ	Attributional Style Questionnaire
AT	As-Treated
BDI	Beck Depression Inventory
BDP	Brief Dynamic Psychotherapy
BT	Behavioral Therapy
CBASP	Cognitive Behavioral Analysis System of Psychotherapy
CBT	Cognitive-Behavioral Therapy
CC	Correlation Coefficient
CDPS	Capacity for Dynamic Process Scale
CT	Cognitive Therapy
DAI	Dynamic Assessment Interview
DAS	Dysfunctional Attitudes Scale
DSM	Diagnostic and Statistical Manual of Mental Disorders
DSQ	Defense Style Questionnaire
ET	Experiential Therapy
FIAT	Flexibility, Interpersonal orientation, Activity, and Teleologic understanding
HARS	Hamilton Anxiety Rating Scale
HDRS	Hamilton Depression Rating Scale
HPS	Helsinki Psychotherapy Study
ICC	Intra-class Correlation Coefficient
IIP	Inventory of Interpersonal Problems
IPT	Interpersonal Therapy
IQ	Intelligence Quotient
ISTDP	Intensive Short-Term Dynamic Psychotherapy
ITT	Intention-To-Treat
LPO	Level of Personality Organization
LTCBT	Long-Term Cognitive-Behavioral Therapy
LTPP	Long-Term Psychodynamic Psychotherapy
MPS	Munich Psychotherapy Study
NEO-PI-R	NEO Personality Inventory – Revised
NML	Nijmegen Motivation List

NR	Not Reported
OCD	Obsessive-Compulsive Disorder
PA	Psychoanalysis
PP	Psychodynamic Psychotherapy
PST	Problem-Solving Therapy
PTSD	Post-Traumatic Stress Disorder
QORS	Quality of Object Relations Scale
RCT	Randomized Clinical Trial
SASB	Structural Analysis of Social Behavior
SAS-SR	Social Adjustment Scale – Self-Report
SCL-90-Anx	Symptom Checklist-90, Anxiety scale
SCL-90-GSI	Symptom Checklist-90, Global Severity Index
SFT	Solution-Focused Therapy
SOC	Sense of Coherence Scale
SPS	Suitability for Psychotherapy Scale
SSCT	Suitability for Short-term Cognitive Therapy (Rating Scale)
STAPP	Short-Term Anxiety-Provoking Psychotherapy
STCT	Short-Term Cognitive Therapy
STPP	Short-Term Psychodynamic Psychotherapy
STS	Systematic Treatment Selection (Clinician Rating Form)
TCI	Temperament and Character Inventory
TORF	Total Object-Relational Functioning
TPV	Total Predictor Variable
URICA	University of Rhode Island Change Assessment Questionnaire
WAI	Working Alliance Inventory
WAIS	Wechsler Adult Intelligence Scale

1 INTRODUCTION

Effective treatment of chronic diseases is the key to an individual's optimal recovery and society's well-being and economy. Due to effective prevention and treatment, the burden of many common chronic diseases has reduced significantly (Knekt et al. 2010). The burden of mental disorders and need for their treatment, however, continues to rise and health care resources have become more and more limited (Knekt et al. 2010). Mental disorders are common in Finland and internationally. One fifth of the Finnish population is estimated to suffer from mental disorders; with mood and anxiety disorders being the most prevalent, at 7% and 4%, respectively (Pirkola et al. 2005). Similar figures have been reported worldwide (Alonso et al. 2004, Steel et al. 2014). Mood and anxiety disorders are widely incapacitating disorders, affecting one's well-being, work ability and functional capacity, social relationships, and quality of life, and they commonly run a recurrent and chronic course (WHO 2000). Work absences and premature retirements related to these disorders, as well as their treatment, cause remarkable costs to society (Sillanpää et al. 2008). Psychotropic medication and psychotherapy are the most common options in the treatment of mood and anxiety disorders (Sihvo et al. 2006, Hämäläinen et al. 2008). Psychotropic medication may help in alleviating the symptoms related to these disorders but not in clarifying and solving the issues underlying and causing them. Psychotherapy aims to achieve this and psychotherapies of different orientation and length are widely applied in the treatment of mood and anxiety disorders.

Suitable psychotherapy is fundamental to recovery from mood and anxiety disorders. The careful selection of patients for psychotherapy has been a crucial part of its development and application ever since Freud first introduced psychoanalysis and set down the criteria for it (Freud 1905a). Psychoanalytic principles served as the basis for the development of analytically oriented long-term psychotherapy and, in the 1940s, Alexander and French (1946) described how the same principles can also be applied to short-term psychotherapy. Although the continuum of psychoanalytic treatments thus expanded into long-term and short-term psychoanalytic therapy, patients were still usually offered longer treatments. From the early 1960s, it was increasingly recognized in outpatient psychiatry clinics that a large percentage of patients were on the waiting list for treatment while a small percentage of patients were being treated with long-term psychotherapy (Davanloo 1978). As the number of treating therapists was limited, it was clear that the only possible solution for responding to the increasing demand for psychotherapy services was a change in the number of patients that each therapist could treat; namely, to increase the number of short-term treatments. Increasing demand for psychotherapies was also increasing the pressure to limit the number of therapy

Patient suitability for short-term
and long-term psychotherapy

sessions covered by health insurance, thus further emphasizing the need for effective short-term treatments (Davanloo 1980). Short-term psychotherapy was not only considered a more economical form of treatment, but it was also argued that a pre-determined, short-term duration could accelerate the pace of work and help to minimize the feeling of dependency on the therapist by patients, thus preventing problems in terminating the treatment. However, the shorter the treatment, the fewer the patients who, in general, were considered suitable and thus treatable with psychotherapy. It thus became essential to separate those suitable for shorter treatments from those in need of longer treatments.

Development and discussion of the criteria needed for the selection of patients suitable for short-term treatment truly began to develop and be applied in the 1960s and 1970s, following the work of Balint et al. (1972), Malan (1963, 1976a), Sifneos (1972, 1979), and Davanloo (1978) on psychodynamic psychotherapies. Based on empirical evidence from a number of patients they had treated with shorter therapy, Balint and Malan, Sifneos, and Davanloo − initially apart and unaware of each other − ended up proposing similar selection criteria for short-term psychodynamic psychotherapy: the establishment of and patient-therapist agreement on a circumscribed psychotherapeutic focus, accessibility and tolerance of affect, ability to interact flexibly with the therapist, realistic self-concept and good ego functions, ability to respond to and elaborate on the therapist's interpretations, capacity for introspection and insight, motivation for treatment, change and psychotherapeutic work, above-average intelligence, flexibility and availability of different defense mechanisms, positive dimensions of personality functioning, and the presence of meaningful relationships in the patient's life. These capacities were considered to enhance the patient's ability to tolerate the anxiety provoked by short-term interpretative psychodynamic psychotherapy, to understand one's inner psychic life, and to commit to working towards solving problems in confided therapeutic alliance with the therapist. These suitability selection criteria, first developed within psychodynamic psychotherapies, were later expanded to other psychotherapy forms and complemented with the psychological capacities considered relevant to them, such as coping styles and cognitive skills in cognitive therapies.

Patient suitability for psychotherapy was considered to be best evaluated through interaction between the patient and an interviewer during an initial assessment period (Davanloo 1978). Such an evaluation should be based on validated psychotherapy suitability assessment scales. Several assessment scales specifically assessing one aspect of patient suitability, such as intelligence (Wechsler 1955), defense mechanisms (Perry 1990, Perry and Høglend 1998), and the quality and quantity of patient's interpersonal relationships (Azim et al. 1991, Benjamin 1996) have been constructed and validated. Scales assessing several aspects of patient suitability have also been constructed; for example Heiberg (1975), based on the work by Sifneos (1972), and Davanloo (1978) presented a scale for the evaluation of

the criteria they proposed for the selection of patients for short-term psychodynamic psychotherapy, and many other psychotherapy suitability assessment scales have later followed them. Only a few of these more comprehensive scales have, however, been studied for their reliability and validity (Heiberg 1975, 1976; Alpher et al. 1990, Baumann et al. 2001, Cromer and Hilsenroth 2010; Safran et al. 1993, Rosenbaum et al. 1997, Jørgensen et al. 2000; Fisher et al. 1999), indicating whether the scale consistently measures what is intended irrespective of an individual interviewer, although a validated instrument is a prerequisite for reliable assessment. The validated scales have concentrated more on some suitability criteria, such as modulation of affects, reflective ability, and motivation, and less on some others, such as self-concept and response to trial interpretation, although these are considered crucial for suitability evaluation (Davanloo 1978). Some of the validated scales also evaluate other aspects of the patients, such as their symptoms (Safran et al. 1993, Fisher et al. 1999). Besides being reliable and valid, a useful psychotherapy suitability assessment scale should also predict the psychotherapy outcome.

The first quantitative, methodologically sound studies investigating the proposed suitability criteria as predictors of psychotherapy outcome emerged in the 1980s (Emmelkamp 1980, Zuckerman et al. 1980, Brodaty et al. 1982, Horowitz et al. 1984, Husby et al. 1985, Husby 1985a, Simons et al. 1985) covering several psychotherapy modalities but with very small sample sizes. However, suitability research increased rapidly during the next two decades, both in numbers of publications and sample sizes. These studies confirmed that, in general, good values in suitability measures served as indications and poor values as contraindications for suitability for the examined therapies. Still, the vast majority of the psychotherapy literature continued to focus on the comparison of the effectiveness of the different forms of short-term psychotherapies that had developed since the 1970s and claimed to challenge psychodynamic psychotherapy in effectiveness. Logically, and according to psychodynamic psychotherapy theory (Malan 1976a), the selection of suitable candidates for therapy should precede the evaluation of the effectiveness of the therapy, as unsuitable patients cannot be expected to benefit from the therapy. Respectively, the comparison of suitability for different psychotherapies should precede the comparison of their effectiveness, as such a comparison is fair only if the patients treated with psychotherapies are considered suitable for and able to benefit from all of them.

Several studies have compared suitability for different types of short-term therapies (Sotsky et al. 1991, Blatt et al. 1995; McBride et al. 2006, Zuroff et al. 2007, Marshall et al. 2008; Joyce et al. 2007, Carter et al. 2011; Rizvi et al. 2009; Wolitzky-Taylor et al. 2012) by comparing the prediction of the outcome of those therapies using the same suitability measures. Although most of these suitability measures were first proposed for short-term psychodynamic psychotherapy, no studies exist comparing their prediction in short-term psychodynamic psychotherapy versus other short-term therapy modalities, and it is therefore unknown whether these

measures work similarly with respect to the therapy for which they were developed and to other therapies. In the only study comparing interpretative and supportive forms of short-term psychodynamic psychotherapies, suitability measures were found to differentiate the outcome of the interpretative form more strongly than that of the supportive form, which is in line with the hypothesis that the supportive form is suitable for a greater variety of patients (Piper et al. 1998). Most comparisons have been performed between cognitive or cognitive-behavioral therapy and interpersonal therapy (Sotsky et al. 1991, Blatt et al. 1995; McBride et al. 2006, Zuroff et al. 2007, Marshall et al. 2008; Joyce et al. 2007, Carter et al. 2011), suggesting some differences in their suitability. The effectiveness of these therapies was, however, evaluated prior to their suitability (Elkin et al. 1989, Imber et al. 1990, Luty et al. 2007). Most studies comparing the effectiveness of short-term therapies thus do so without first giving consideration to whether they are equally suitable.

Psychotherapy suitability research has almost entirely focused on short-term psychotherapies; only a few suitability studies on long-term psychotherapies have been carried out (Jørgensen et al. 2000, Puschner et al. 2004, Solbakken et al. 2012). Most importantly, no studies comparing suitability for short-term and long-term therapy have been published, even though suitability criteria were developed above all in order to separate those suitable for short-term therapy from those apparently needing long-term therapy in order to recover. For short-term therapy to be truly suitable for a given patient it should afford equal or greater benefit for that patient than long-term therapy. Similarly, for long-term therapy to be suitable for patients unsuitable for short-term therapy, it should be shown to bring them significant benefits. Evidence-based information on patient suitability for short-term and long-term psychotherapy is thus urgently needed.

In this study, patient suitability for short-term and long-term psychotherapy is compared for the first time. Patients suffering from mood and anxiety disorders, considered treatable with both short-term and long-term psychotherapies, are included. A new interview-based psychotherapy suitability assessment scale, the Suitability for Psychotherapy Scale (SPS) − evaluating focus, modulation of affects, flexibility of interaction, self-concept in relation to ego ideal, response to trial interpretation, reflective ability, and motivation − is introduced and analyzed for its reliability and validity. Prediction of the outcome by the SPS in short-term psychodynamic psychotherapy and another short-term therapy − solution-focused therapy − as well as in these short-term therapies and in long-term psychodynamic psychotherapy is compared for the first time.

2 REVIEW OF LITERATURE

2.1 Development of psychotherapies

At the turn of the 20th century, Sigmund Freud, with the help of Josef Breuer, developed the classical psychoanalytic technique and psychoanalysis (Breuer and Freud 1895, Freud 1905a) which later formed the basis for the development of long-term psychoanalytic psychotherapy and short-term psychoanalytic psychotherapy. Freud emphasized the understanding of human behavior from a psychodynamic point of view, and investigation of the psychological processes underlying symptoms. Psychoanalysis thus involved looking at how the unconscious mind affects thoughts, feelings, and behaviors and how early childhood experiences contribute to current actions. Psychoanalysis became the standard treatment offered to patients suffering from a wide range of neurotic and character problems. Although Freud himself performed and published brief analyses (Freud 1895, 1905b, 1909), the evolvement of the psychoanalytic technique from hypnosis through suggestion to free association transformed psychoanalysis into an ever lengthier and more complex therapeutic procedure (Malan 1963, Balint et al. 1972). At the same time, it shaped the role of the therapist as increasingly passive. The dominance of psychoanalysis and its supporters made it difficult for anyone to try to develop a shorter form of psychoanalytic treatment emphasizing a more active role for the therapist. Only seven published reports of such shorter treatments, carried out by analysts other than Freud, can be found for the period between 1909 and 1920, and even fewer of these are accompanied by a follow-up showing whether sustained improvement was attained (Malan 1963).

During the 1920s, following the tremendous initial enthusiasm for psychoanalysis, growing opposition towards its increasing length and the passivity of the therapist was, however, starting to develop (Malan 1963). During and after the First World War the need to find shorter forms of psychoanalytic treatment was also recognized (Balint et al. 1972). Freud (1919) already accurately anticipated society's needs and demands for a treatment available for a larger segment of the population, free of charge, and implied that the 'pure gold of analysis' might need to be replaced by the 'copper of direct suggestion' in order to respond to that demand (Balint et al. 1972). The birth of psychoanalytic ego psychology, initiated by Freud (1923), gave rise to the development of a wide range of dynamically oriented psychotherapeutic approaches, which emphasized the psyche's both conscious and unconscious influence on the self and external relationships. Theorists such as Anna Freud, Alfred Adler, Carl Jung, Karen Horney, Erik Erikson, and Melanie Klein, building upon but departing from Freud's fundamental ideas, developed their own schools of psychotherapy. It was, however, Sándor Ferenczi and Otto Rank who

foremostly took up the challenge and experimented with new shorter, more active therapeutic approaches (Ferenczi 1919, 1920, 1925, Ferenczi and Rank 1925). However, their critical review of and modifications to the psychoanalytic therapeutic process were widely opposed by Freud and others, as a result of which the development of other forms of shorter psychotherapy was inhibited for some time (Balint et al. 1972).

Demand for shorter psychotherapies resumed during and after the Second World War and numerous analysts engaged with the question of how to shorten the course of therapy but still achieve therapeutic effectiveness (Balint et al. 1972). Under the leadership of Franz Alexander and Thomas Morton French, the Chicago Psychoanalytic Institute began new, more systematic experiments on more active, short psychoanalytic psychotherapies in 1938. Based on this work, they were able to show that it was possible to achieve not only symptomatic change, but even characterological change, through less intensive treatments, thus demonstrating the effectiveness and value of short-term therapy (Alexander and French 1946). Alexander (1925) had critiqued Ferenczi and Rank for making overly substantial changes to psychoanalytic techniques; although Alexander and French (1946) presented their work within a psychoanalytically oriented framework, it still represented a modification of, rather than a short-term alternative based on, psychoanalytic principles (Malan 1976a). Thus, most psychoanalysts could not accept that this form of psychotherapy was on a continuum with psychoanalysis. In particular, the recommendations by Alexander and French (1946) for changes in the therapist technique, turning the therapist from a passive listener into an active participant observer, initiated a heated dispute and finally led to a demarcation between psychoanalysis and psychoanalytic psychotherapy (Balint et al. 1972). Although the attempts of Alexander and French to develop short-term psychotherapy faced a great deal of opposition at the time, their eventual impact was tremendous as it led to various other attempts, met by steadily growing interest, and more lasting efforts to develop shorter forms of psychotherapy.

Following Alexander and French (1946), various other models of short-term psychodynamic psychotherapy were proposed, for example by Baker (1947), Fenichel (1954), Socarides (1954), Burdon (1963), Bellak and Small (1965), Gillman (1965), McGuire (1965), Wolberg (1965), and Mann (1973). The aim of all these models to make therapy as brief as possible separated them from psychoanalysis and long-term, open-ended psychotherapies. According to Bellak and Small (1965), symptomatic change could ideally occur in as few as five sessions of customary duration (45 to 60 minutes). Mann (1973) was, however, the only therapist to set an absolute time-limit of 12 sessions; he considered a limited number of sessions to be a curative factor in itself, enabling greater consistency in the study of the meaning of time in short-term psychotherapy. The pioneers in the systematic, research-based development of well-defined short-term psychodynamic psychotherapies in the 1950s-1970s were, however, Michael Balint and David

Malan at the Tavistock Clinic in London, Peter Sifneos at the Massachusetts General Hospital in Boston, and Habib Davanloo in the General Hospital and McGill University in Montreal (Malan 1963, Balint et al. 1972; Sifneos 1965, 1972; Davanloo 1978, 1980).

Balint began case discussion seminars for general practitioners in 1952 and workshops on brief psychoanalytically oriented psychotherapy for a small team of highly qualified psychotherapists in 1954, with Malan as a founding member (Malan 1976a). Experience gained via clinical practice supported the theoretical suggestion that active, interpretative short-term treatments could bring about lasting symptomatic and characterological changes even in relatively severe and chronic conditions in a maximum of 40 sessions. Based on this work, Balint, Malan, and their colleagues developed a brief form of psychotherapy called 'focal psychotherapy', in which one specific problem presented by the patient is chosen as the focus of interpretation (Balint et al. 1972). By repeatedly interpreting a) the patient's defenses, used to manage anxiety, caused by hidden feelings rising into awareness, by pushing them back into the unconscious (triangle of conflict), and b) the transference to the therapist, reflecting the maladaptive patterns in the patient's current relationships, triggered by the patient's past and early childhood experiences with the parents, which generated the hidden feelings (triangle of person), the patient is believed to acquire insight into these basic conflicts and to be able to modify his or her defensive strategies. This model of clarifying and working through the patient's resistance using these two triangles, originally proposed by Ezriel (1952) and Menninger (1958) and further adapted and expanded by Malan, differed in its active nature from the traditional psychoanalytic technique, in which, according to Malan, the biggest mistake was to react to increasing resistance with increased passivity (Malan 1986). Malan recommended a shorter time limit of 30 sessions, or even 20 for an experienced therapist (Malan 1976a). Balint and Malan still considered focal psychotherapy to be on a continuum with psychoanalysis, but due to the number of modifications to traditional psychoanalytic techniques it continued to face resistance and critique. In Balint's own words "it was very difficult indeed to realize that the new techniques and way of thinking did not endanger basic psychoanalytic theory and practice: that they were supplementary and not antagonistic to each other" (Balint et al. 1972).

Concurrently with Balint and Malan, Sifneos was developing a similar, even a slightly more radical psychotherapy than focal psychotherapy, which he termed Short-Term Anxiety-Provoking Psychotherapy (STAPP), with an upper limit of 20 sessions, as opposed to his largely supportive anxiety-suppressive psychotherapy, which was usually of a longer duration (Sifneos 1972). STAPP deliberately raises the patient's anxiety from early on by concentrating on interpreting defenses and resistance expressed through transference (Sifneos 1966, 1967, 1972), in a similar manner to focal psychotherapy. However, STAPP differs from focal psychotherapy by further emphasizing the learning of new ways of solving emotional problems as a

criterion of a successful outcome; a resource that may be utilised long after treatment has terminated in order to prevent a relapse or the reoccurrence of the problems.

The research groups lead by Balint and Sifneos only became aware of each other's work when they met at the Sixth International Congress of Psychotherapy in London in 1964 (Malan 1976a). Around the same time, Habib Davanloo had initiated a weekly seminar for the residents and medical students of the Montreal General Hospital and began systematic research into a special kind of short-term psychoanalytic psychotherapy which he called Intensive Short-Term Dynamic Psychotherapy (ISTDP; Davanloo 1978). This involved the same essential therapeutic elements as focal psychotherapy and STAPP and an active and persistent confrontation of the patient's resistances similar to that of STAPP, and typically lasted 20-40 sessions. From 1963 to 1974, Davanloo evaluated 575 patients, about 40 percent of whom he was able to follow for 2-7 years. In 1973, while attending the Ninth International Congress of Psychotherapy, Davanloo met Sifneos and became aware of his and Malan's work. In the following year, Davanloo invited Sifneos to collaborate with him in a Montreal symposium on short-term dynamic psychotherapy, and visited Malan and attended his workshop in London. In 1975 and 1976, Davanloo set up the First and the Second International Symposium and Workshop on Short-Term Dynamic Psychotherapy in Montreal, which primarily focused on the work by Davanloo, Sifneos, and Malan, with 600 mental-health professionals from all over the world attending. Their work was also brought together in several multi-authored textbooks (Davanloo 1978, 1980), which had a major impact on the field of short-term psychodynamic psychotherapy. The fact that three research groups came to such similar conclusions independently of each other supported the likelihood of the existence of a common psychoanalytic short-term psychotherapy technique, on which careful research could converge.

In the 1950s-1970s, several other forms of psychotherapies also emerged to challenge the dominant paradigm, such as behavioral, cognitive, cognitive-behavioral, interpersonal, problem-solving, solution-focused, and experiential therapies. The roots of behavior therapy lie in the 1950s when B. F. Skinner and his colleagues in the US, inspired by the classical conditioning theory of Ivan Pavlov and behaviorist learning theory by John B. Watson, suggested the use of operant conditioning, which involves the use of rewards and punishments to strengthen or weaken certain behaviors, to improve the functioning of people with mental health disorders, and anxiety disorders in particular (Lindsley et al. 1953). In 1952 in the UK, Hans Eysenck first introduced the term behavior therapy (BT) in reference to a type of treatment that aims to modify maladaptive behavior by learning, either by personal exposure or via modeling, more adaptive ways of behaving, as opposed to exploring the underlying causes behind the maladaptive behavior in psychodynamic psychotherapy (Yates 1970). In the 1960s, Aaron T. Beck, together with his colleagues at the University of Pennsylvania in Philadelphia, developed cognitive

therapy (CT), which seeks to help patients suffering from depression and other mental health disorders by identifying, reality-testing, and correcting the distorted cognitions or patterns of thinking that are believed to underlie the disorders (Beck 1970, 1976, Beck et al. 1979). Whereas BT tends to be very short, consisting of just a few sessions, CT can be either short, typically 12-20 sessions, or long and open-ended. The joining of the techniques of BT and CT – i.e., modification of behaviors in combination with the beliefs that are believed to be causing them – gave rise to cognitive-behavioral therapy (CBT; Rachman 1997). A typical CBT lasts 10-20 sessions. The duration and frequency of CBT, as well as BT and CT, should be adjusted according to the type and severity of the disorder and individual needs. Interpersonal therapy (IPT) was developed at the turn of the 1960s and 1970s by Gerald Klerman, Myrna Weissman, Albert DiMascio, and their colleagues in the New Haven-Boston Collaborative Depression Project for the treatment of ambulatory nonpsychotic, nonbipolar depression (Klerman et al. 1979, 1984). IPT uses techniques derived from psychodynamically oriented therapies but focuses on the patient's current life and interpersonal problems, rather than on early developmental experiences. The goal of IPT is to develop more adaptive ways of relating to others during a relative short time frame (12-20 sessions), and it is believed that improvements in other areas of life will follow improvements in interpersonal functioning. Problem-solving therapy (PST) developed around the same time as IPT, out of a trend which involved teaching patients problem solving and psychosocial skills during psychotherapy of very short duration, typically 4-8 sessions (D'Zurilla and Goldfried 1971, D'Zurilla and Nezu 1999). The goal of PST is for the patients to learn how to use and reactivate their own skills to solve both present and future problems in a step-by-step process. Due to its short duration and practical approach, PST has also been considered suitable for use in primary care by general practitioners. Solution-focused therapy (SFT) is another very short psychotherapy approach, with a typical upper limit of 12 sessions, which was developed by Steve de Shazer, Insoo Kim Berg, and their colleagues (de Shazer et al. 1986, de Shazer 1991) in the late 1970s and early 1980s in the US. Similarly to IPT and PST, SFT focuses on the present and, in particular, the future rather than the past, but differs from PST in its focus on resources and solution-building rather than problems and problem-solving. Experiential therapy (ET) is another here-and-now oriented psychotherapy which was developed by Alvin Mahrer in the 1970s and has its roots in existential philosophy and Gestalt therapy (Mahrer 1983, 1989, 1996). ET may, however, be of short or long duration, depending on individual needs. In ET, each person is seen as an individual who applies meaning to his or her external world and relationships, and has potential for personal growth and self-transformation, which is also the goal of the therapy. Mutual involvement of the therapist and the patient is emphasized and the patient is invited to actively participate, for example in role-playing via an empty-chair dialogue. The client-centered therapy, currently referred to as person-centered therapy and developed

based on the writings of Carl Rogers (1951, 1957, 1961), emphasizes the importance of therapist empathy, authenticity, and congruence. It is considered a form of experiential therapy. As in cognitive-behavioral therapy, in which elements of two therapies were combined, more therapies building upon elements of various others were developed later, with the purpose of more adequately modeling the elements of human complexity. The latest developments include acceptance and commitment therapy (ACT) developed for the treatment of anxiety disorders, with its roots in the behavioral and experiential therapy traditions (Hayes et al. 1999), and the cognitive behavioral analysis system of psychotherapy (CBASP) developed for the treatment of chronic depression, building upon cognitive-behavioral, interpersonal, and also partly on problem-solving therapy (McCullough 2000). ACT cultivates acceptance of experiencing actively and fully in the present, the exploration of personal values, and commitment to behavior change aligned with those personal values (Hayes et al. 1999). CBASP helps chronically depressed patients to understand the consequences of their behavior for their environment and interpersonal relationships, and aims to improve their interpersonal skills by teaching them social problem-solving skills and alternative ways of coping (McCullough 2000). These and other therapy forms are constantly evolving, leading to the modification and integration of theory and techniques, thus further contributing to the manifold field of psychotherapy.

When any new psychotherapy model is introduced, in addition to describing how the model addresses the problems of the patient and the techniques used to achieve this, two other important issues should also be addressed: 1) definition of selection criteria, i.e., which patients are considered suitable for the psychotherapy, and 2) effectiveness of the psychotherapy, i.e., what kind of therapeutic results can be achieved and how they compare with other treatments available for the same target population (Malan 1963). In the following, both the suitability for (Section 2.2) and the effectiveness of (Section 2.3) psychotherapies of different orientation and length are discussed.

2.2 Suitability for psychotherapies

When a patient first seeks treatment for psychiatric problems, an initial clinical evaluation should be performed to 1) determine the reasons for the patient's suffering, and 2) to choose an optimal treatment for the patient. The choice of optimal treatment requires the definition of selection criteria for psychotherapies of different length and orientation. Selection of patients for a particular treatment covers both the ruling out of certain contraindications, i.e., the elimination of obviously unsuitable patients, and the determination of certain indications, i.e., the choice of likely suitable patients (Malan 1976a). One relevant aspect in the evaluation of the patient's suitability is the type, severity, and duration of the patient's disorder, as different psychotherapies are intended for the treatment of different types of disorders and longer psychotherapies are generally considered

necessary for the treatment of more severe and chronic disorders (see Section 2.1). Accordingly, the initial evaluation typically begins with an assessment of the patient's psychiatric history and symptoms, and the formulation of a psychiatric diagnosis. These alone do not, however, determine the patient's psychotherapeutic needs and potential; patients with the same diagnosis or identical symptoms may reveal different symptom origins and present with varying capacities for psychotherapeutic work. To evaluate these capacities prior to the start of therapy, it has been suggested that a kind of trial therapy should be carried out as part of the initial evaluation, exposing the patient to some of the important elements and techniques of the therapy, such as questioning, confrontation, resistance, working through, and transference in psychodynamic psychotherapies (Davanloo 1978). Each intervention would be met with a response from the patient, which would in turn provide evidence on his or her psychological capacities, considered of primary importance in psychotherapeutic work and thus to be another important aspect and major determinant of psychotherapy suitability and the related outcome.

Selection of patients suitable for a particular treatment has been emphasized since the early development of psychoanalysis. Freud (1905a) outlined four major criteria that the patient must fulfill to be considered analyzable and to undergo psychoanalysis: the patient must 1) suffer from a chronic neurotic syndrome ('transference neurosis'), 2) have passed through adolescence but not yet reached fifty years of age, 3) be of high intelligence, and 4) be of reliable character. These criteria thus covered both the aspects of the disorder and the patient's characteristics. Patients not fulfilling these criteria were considered to suffer from narcissistic neuroses or lack the necessary therapeutic resources, therefore being untreatable with psychoanalysis. At first, when no or only little empirical data was available, indicators of treatment suitability were necessarily based on preconceived ideas which were gradually sharpened as empirical evidence kept accumulating. As a result, updated, comprehensive guidelines on contraindications and indications for psychoanalysis are nowadays available (American Psychiatric Association 1985, Knekt et al. 2011a). Regarding psychiatric diagnoses, psychotic disorders and severe personality disorders are considered contraindications for psychoanalysis, whereas patients with mainly oedipal (neurotic) core conflicts and chronic symptoms, reflecting intrapsychic conflict and developmental arrest, are considered treatable with psychoanalysis. Very severe early trauma and a severe current life crisis are, however, considered contraindications for psychoanalysis. A balance between a sufficient amount of subjective suffering and sufficient potential for growth is considered an essential indication for psychoanalysis. The patient characteristics considered to reflect such growth potential comprise sufficient ego strength (i.e., the ability to tolerate anxiety and therapeutic regression, control impulses, test reality, and interact flexibly), psychological mindedness (i.e., motivation for self-exploration and the ability to make psychological connections and work with trial interpretation), and an adequate developmental level of object relations. In general, a

need for psychoanalysis is usually indicated when response to other psychiatric treatments is expected to be inadequate.

The criteria for patient suitability for long-term and short-term psychotherapies have developed hand in hand since short-term psychotherapies began to emerge (Section 2.1) with the aim of developing the capability to select and separate patients suitable for short-term psychotherapy from those in apparent need of long-term psychotherapy. Development of the suitability criteria was thus mainly carried out within short-term psychotherapies, psychodynamic psychotherapies in particular, in order to demonstrate their utility and applicability. Suitability criteria for short-term psychotherapy have therefore mainly been discussed and studied, it being implicitly assumed that patients not fulfilling these criteria, and thus not suitable for short-term psychotherapy, are suitable for and able to benefit from long-term psychotherapy.

Views on the diagnostic indications for short-term psychotherapy differed radically at first. According to a conservative, 'static' view, short-term psychotherapy was only to be used when long-term psychotherapy was not available, and was considered suitable only for mild illnesses of acute and recent onset (Knight 1937, Fuerst 1938, Berliner 1941), i.e., for patients "who are not very sick in the first place" (Barron 1953, Sullivan et al. 1958). Supporters of this view recommended that the techniques used should be superficial and avoid transference interpretation in particular. They held the belief that short-term psychotherapies essentially offered temporary, palliative symptom relief. According to a radical, 'dynamic' view, on the other hand, a more extensive and long-standing character pathology was considered treatable with interpretative short-term psychodynamic psychotherapy, if the patient showed a willingness and ability to work within such a framework (Malan 1963). Alexander and French (1946), advocates of the radical view, demonstrated the broad effectiveness of short-term psychodynamic psychotherapies (see Section 2.1). Empirical work by Balint's group (Malan 1963), based on treating patients, who presented with more severe and chronic illnesses, with interpretative short-term psychodynamic psychotherapy also strongly supported the radical view (see Section 2.1). Furthermore, they found interpretation of transference throughout the therapy to be the most important technique in terms of achieving a successful outcome, the opposite to what was believed by supporters of the conservative view. Their findings also supported substantial, deep-seated improvements at the end of the psychotherapy which persisted, and in some patients even increased, during a long post-treatment follow-up, thus also contradicting the conservative views on the superficiality and unsustainability of the short-term psychotherapy results. Whereas supporters of the conservative view mainly based their opinions on clinical impressions and experiences, supporters of the radical view emphasized the importance of empirical research of theoretical or clinical proposals and based their conclusions on that. As the empirical evidence kept growing, the general view began to shift towards the radical end of the spectrum (Malan 1976a).

Regarding the indications related to patients' psychological capacities, supporters of both the conservative view (Ripley et al. 1948, Barron 1953, Pumpian-Mindlin 1953, Rosenbaum et al. 1956) and of the radical view (Alexander 1944) stressed the paramount importance of a well-adjusted personality and high ego strength, judged by the patient's ability to bear frustration and conflict, cope with reality, and form satisfactory interpersonal relations, to a good prognosis in short-term psychotherapy. It was, however, the supporters of the radical view who developed and tested a more refined and comprehensive set of psychological suitability selection criteria, based on such intrapsychic and interpersonal capacities, systematically over time. This development began concomitantly with the development of short-term psychodynamic psychotherapies and was led by its pioneers: Balint, Malan, Sifneos, and Davanloo (see Section 2.1).

The selection criteria initially suggested for short-term focal psychotherapy, when the workshops run by Balint and Malan at the Tavistock Clinic in London began in 1954 (see Section 2.1), were largely based on preconceived ideas that a young age and fairly reliable ego-structure, demonstrated by life achievements and the ability to form lasting object-relationships, were important to a successful treatment outcome (Malan 1963). As the workshops progressed and more insight was gained, these criteria gradually developed into more detailed and reliable criteria, tested against the successful psychotherapy outcome: 1) the evaluator's ability to understand the patient's problem in dynamic terms, 2) the evaluator's ability to formulate some kind of circumscribed therapeutic plan, 3) the patient's willingness and ability to explore feelings, and 4) the patient's ability to work within a therapeutic relationship based on interpretation (Malan 1963). It was recommended that those who seemed suitable after the first interview should be interviewed again a week later to allow the observation of the patient's likely reactions to interpretative psychotherapy over a longer period. Later Balint further emphasized the clinical impression of the patient's motivation to change and ability to eventually develop a reliable therapeutic alliance (Balint et al. 1972). He also outlined that the focus should be found no later than by the third or fourth session. Balint did not, however, provide any statistical evidence to support the validity of these criteria. Finally, after two decades of continuous work with two large patient samples, Malan (1976a) described the following steps in the selection of patients for focal psychotherapy. First came the exclusion of patients with absolute contraindications. An unpublished but at the time generally accepted list of contraindications for short-term psychodynamic psychotherapy was provided by Hildebrand at the London Clinic of Psychoanalysis (Malan 1963, 1976a): serious suicidal attempts, gross destructive or self-destructive acting out, serious drug addiction or alcoholism, convinced homosexuality, long-term psychiatric hospitalization, several courses of electroconvulsive therapy, and incapacitating chronic obsessional or phobic symptoms. Second, further rejection of patients in whose case certain dangers seem inevitable. These dangers include a) inability to

begin and carry out effective therapeutic work within a short timeframe due to inability to make contact, lack of sufficient motivation, or the presence of rigid defenses; b) inability to terminate therapy within the short timeframe due to the involvement of issues that are too complex and deep-seated, severe dependence or other forms of unfavorable intense transference; and c) breakdown due to the intensification of a depressive or psychotic disturbance, suicide, or uncontrollable acting out. For such patients, prolonged work in long-term psychotherapy was considered necessary to generating motivation, penetrating rigid defenses and thoroughly working through and resolving complex issues. Third, for a patient to be considered a suitable candidate, the evaluator should be able to formulate a focus, including the identification of a) the 'nuclear' conflict, b) the current conflict, and c) the congruence between the nuclear and current conflicts, i.e., whether they can be viewed as being essentially the same. Fourth, this focus should be acceptable to the patient, who is therefore required to demonstrate a) good capacity to consider his or her problems in emotional terms, b) sufficient strength to face disturbing material, c) a good response to interpretations relevant to the focus which convincingly confirms them, and d) motivation for insight. Based on the research evidence, Malan concluded that motivation for insight correlates most strongly with favorable treatment outcome, and is thus the most important selection criterion. Malan, however, also emphasized that these suitability criteria were ideal, not absolute, and that the balance between the criteria should also be taken into account in the suitability evaluation.

While Balint and Malan were developing the suitability selection criteria for their short-term focal psychotherapy in London, Sifneos was developing similar criteria for his Short-Term Anxiety-Provoking Psychotherapy (STAPP) at the psychiatry clinic of Massachusetts General Hospital in Boston (Sifneos 1965, 1968a, 1972, 1979). Similarly to Balint and Malan, he also held a continuous seminar involving case presentations of patients undergoing STAPP, which served as the basis for the formulation of the suitability criteria (see Section 2.1). Since STAPP focuses on the resolution of the neurotic conflicts that underlie the patient's symptoms, Sifneos initially emphasized the emergence of a circumscribed neurotic problem or focus (i.e., nuclear conflict) during the first interview, as a necessary condition for selection for STAPP (Sifneos 1965). Additional criteria for an ideal STAPP candidate were: 1) a specific chief complaint, referring to the patient's ability to prioritize one of his or her complaints over others (i.e., current conflict), 2) ability to express feelings and interact flexibly with the evaluator, 3) motivation to work hard during psychotherapy, 4) above-average intelligence, 5) at least one meaningful relationship with another person during the patient's lifetime, and 6) an emotional crisis (Sifneos 1965). If the patient fulfilled at least three of these six criteria, he or she was offered a second interview which served a dual purpose: a) to ensure that the therapist and the patient agreed on the aim of the therapy, and b) to more thoroughly assess the patient's motivation, not only for symptom relief but also for

change, which Sifneos considered the most important suitability criterion. The evaluation of motivation comprised the following seven sub-criteria: 1) the ability to recognize that the symptoms are psychological in nature, 2) a tendency to be introspective and to give an honest and truthful account of emotional difficulties, 3) willingness to participate actively in the treatment situation, 4) curiosity and willingness to understand oneself, 5) willingness to change, explore, and experiment, 6) realistic expectations of the results of psychotherapy, and 7) willingness to make reasonable sacrifices in terms of time and fees (Sifneos 1968a). Patients who fulfilled all these seven criteria were considered to have excellent motivation, patients who fulfilled six, five, or four criteria were considered to have good, fair, and questionable motivation, while patients who fulfilled three criteria or less were considered unmotivated. In their later, updated psychotherapy suitability guidelines (Sifneos 1972), Sifneos and his colleagues omitted the criterion regarding the emotional crisis, but kept the guidelines for the assessment of motivation the same. They also allowed for the establishment of therapeutic focus, in agreement with the patient, to follow the selection of a suitable patient for STAPP.

The suitability selection criteria developed by Sifneos (1972, 1979) are generally considered the most specific and demanding set of suitability criteria (Barth et al. 1988a). Barth et al. (1988a) have suggested a generalization of the initial assessment form for STAPP, applicable also to the selection of patients for short-term focal psychotherapy developed by Balint and Malan, in order to address the extensive overlap of the suggested suitability criteria. According to this generalization, patients fulfilling 4-5 main criteria (circumscribed chief problem, circumscribed chief complaint, emotional interaction with evaluator, above-average problem-solving capacity, and history of meaningful relationship) and 5-7 motivation criteria are accepted for STAPP (20 sessions), whereas patients fulfilling only 2-3 main criteria, but still 5-7 motivation criteria, are accepted for short-term focal psychotherapy (40 sessions). Barth et al. (1988a) further suggested that the rest of the patients could be offered a so-called FIAT (Flexibility, Interpersonal orientation, Activity, and Teleologic understanding) approach (Nielsen et al. 1984), which is an integrated form of short-term psychodynamic psychotherapy, involving the application of supportive, behavioral, and cognitive procedures. More supportive forms of short-term psychodynamic psychotherapy are generally considered suitable for a wider range of patients than interpretative forms of short-term psychodynamic psychotherapy.

The third most comprehensive set of psychotherapy suitability selection criteria has been developed by Davanloo (1978) for his Intensive Short-Term Dynamic Psychotherapy (ISTDP). Davanloo emphasized the importance of evaluating the patient's state of shame and guilt, the level of aggression, and the way the patient deals with this, in order to evaluate the presence of self-destructive impulses and behavior, and thus the contraindication of suicidal risk (Davanloo 1978). Later, he proposed an updated list of contraindications for ISTDP: serious suicidal attempts, a

history of previous psychotic decompensation, paranoid conditions, and poor impulse control (Davanloo 1986). In line with Balint and Malan, and Sifneos, Davanloo too considered the establishment of a psychotherapeutic focus of the treatment, in agreement between the evaluator and the patient, to be of major importance. Whereas Sifneos (1972) and Malan (1976a) emphasized the selection of patients with oedipal chief problem, i.e., stemming from the oedipal period, and thus oedipal focus, Davanloo (1978) considered both patients with an oedipal and non-oedipal focus to be suitable for ISTDP. However, the more the central core of the patient's problem was fixated at the non-oedipal, pre-genital level, the more problematic the therapeutic course was expected to be, for example due to dependency problems. In addition to the evaluation of focus, Davanloo considered the evaluation of seven different aspects of ego functions essential to determining whether the patient had sufficient indications for ISTDP: 1) the affective function of the ego, 2) the response to interpretation, 3) psychological mindedness, 4) motivation, 5) intelligence, 6) the ego's defensive psychological organization, and 7) the quality of human relationships. Of all these criteria, Davanloo considered a good response to the trial interpretation to be the most important.

In psychotherapy forms other than psychodynamic psychotherapy, much less emphasis has been placed on the systematic evaluation of patient suitability, although sufficient ego strength for facing the uncovering psychotherapy and motivation for change are generally considered essential to any form of psychotherapy. In behavior therapy, a clear focus of the treatment, typically the single-most troubling problem, is considered essential but little attention is paid to other determinants of psychotherapy suitability (Strupp and Binder 1984). In cognitive therapy, in addition to cognitive skills, the selection criteria have traditionally focused on a patient's diagnosis, psychopathology, and response to homework assignments (Beck et al. 1979, Fennell and Teasdale 1987, Persons et al. 1988). Later, more emphasis has been laid on the role of emotion, defensive information processing, interpersonal relationships, and the therapeutic relationship within the psychotherapeutic process (Safran et al. 1993). Consequently, similar selection criteria to short-term psychodynamic psychotherapies, based on the above-described work of Balint and Malan (Malan 1976a), Sifneos (1972), and Davanloo (1978) as well as the work of Bordin (1979) on therapeutic alliance, have also been proposed and tested in cognitive therapy with promising results (Safran et al. 1993). The criteria originally derived from psychodynamic psychotherapeutic practice have also been adapted and applied in cognitive-behavioral therapies (Blenkiron 1999). In interpersonal therapy, a specific interpersonal focus for treatment, a relatively secure attachment style, ability to relate to the therapist and others, and a good social support system are thought to increase the patient's suitability for therapy (Stuart and Robertson 2003). It is also considered critical that the patient agrees with such intervention and finds it to be in accordance with his or her psychological problems and their origin.

An initial evaluation of both the reasons for the patient's suffering and the patient's suitability for particular treatment is generally considered to require more than one interview. According to both Malan (1976a) and Sifneos (1972), the final decision regarding a patient's suitability for focal psychotherapy or STAPP can be made in two interviews. The second interview is considered crucial, especially for the thorough evaluation of certain aspects of patient suitability, such as affective functioning, the response to the trial interpretation, and motivation. Also, Davanloo (1978) considered a second or even third interview necessary to allowing the motivation to develop. More generally, according to most short-term psychodynamic therapists, the relevant information regarding the assessment of psychotherapy suitability can be obtained in one to three interviews (Strupp and Binder 1984), and some therapists, such as Mann (1973), allow up to four evaluative sessions. Advocates of other psychotherapy forms, however, seem to place less emphasis not only on the evaluation of suitability but also on its thoroughness, often suggesting only one initial interview (Safran et al. 1993, Blenkiron 1999). According to Malan (1976a), accurate patient selection is "probably the most complex, subtle, and highly skilled procedure in the whole field". Thorough training in the evaluation of psychotherapy suitability is therefore also needed.

Based on the overall research evidence on short-term psychotherapies and their suitability, the following six conclusions can be made. First, impressive psychotherapeutic results can be achieved with a limited number of therapy sessions, and these may be neither superficial nor temporary but rather profound and lasting. Second, short-term psychotherapy is not only effective in the case of mild neurotic conditions of recent onset but good results can also be obtained among those suffering from neurotic or characterological disorders of many years' duration. Third, the major objective of the initial assessment is not just to arrive at a clinical diagnosis but also to determine the focus of and suitability for psychotherapy. Fourth, to determine psychotherapy suitability, a focused interview, which includes trial interventions and thus resembles the therapeutic process, should be carried out. Fifth, the suggested psychotherapy suitability selection criteria vary between advocates of both the same and different psychotherapy orientations, but have considerable overlaps. Sixth, the range of patients considered treatable with short-term psychotherapies has been progressively extended. Nowadays, the generally accepted diagnostic contraindications for short-term psychotherapies include evidence of severe psychopathology, such as psychosis, severe depression, severe impulse control disorder, active alcohol or drug dependency and abuse, organic illnesses, and multiple long-term disorders (Blenkiron 1999). To identify patients likely to be treatable with short-term psychotherapy, the following list of 12 indications for short-term psychotherapy can be presented: 1) patient's ability to identify and define key problems based on which a well-circumscribed focus of treatment can be established, in agreement between the interviewer and the patient; 2) a high degree, accessibility, and tolerance of affect; 3) ability to engage in flexible

and emotional interaction with the interviewer; 4) realistic self-concept and good ego functions; 5) ability to respond to and elaborate on interpretations relevant to psychotherapeutic focus; 6) good capacity for introspection, insight and verbal communication of one's thoughts, feelings, and inner psychic life; 7) genuine conscious and unconscious motivation for treatment, insight, and change; 8) above-average intelligence or problem-solving capacity; 9) maturity, flexibility, and the availability of defense mechanisms; 10) adaptive coping styles and good cognitive skills; 11) positive dimensions of personality functioning; and 12) the ability to form relationships, reflected by a history of mutual emotional involvement with significant others. Such resources are considered to reflect a flexible and healthy personality, necessary in order for a patient to be able to invest in and gain from a psychotherapeutic relationship of shorter duration. Conversely, the absence of such resources is considered to contraindicate the sufficiency of short-term psychotherapy and indicate a need for long-term psychotherapy in order to achieve a sustained recovery. Accordingly, such psychological capacities are proposed as psychotherapy suitability selection criteria, capable of separating patients likely to be treatable with short-term psychotherapy from patients apparently in need of long-term psychotherapy (Table 1). However, it has also been suggested that these different psychological capacities affect and possibly balance each other (Malan 1976a, Davanloo 1978). Evaluation of overall suitability is therefore considered to be of the utmost importance (Table 1).

2.3 Effectiveness of psychotherapies

Psychoanalysis (PA) is the oldest therapeutic discipline with a strong theoretical background, but empirical evidence for the effectiveness of PA is scarce. An apparent reason for this is the length and frequency of PA, typically comprising two to five sessions a week over three to seven years (de Jonghe et al. 2012, de Maat et al. 2013), which make it a very time-consuming and expensive object of study. In the hierarchy of evidence-based medicine, randomized clinical trials (RCT), in which patients are randomly assigned to two or more groups, ideally including the treatment of interest and a non-treatment or waiting-list control condition, are generally considered the gold standard of effectiveness – or, in ideal circumstances, efficacy – research (American Psychiatric Association 1993, Seligman 1995, GRADE Working Group et al. 2004). Unless the treatment is controlled for, it is not possible to evaluate the amount of change in the severity of a mental disorder over time due to the treatment, and not due to the expectation of improvement or regression of the symptoms (which are typically at their greatest when seeking treatment) towards the mean (Bland and Altman 1994, Linden 2013). Such a control condition cannot, however, be included in the effectiveness studies of long- term treatments due to non-feasibility (i.e., many patients are unlikely to accept such an assignment, thereby making the study population non-representative) and ethical

Table 1. Psychotherapy suitability selection criteria based on patient's psychological capacities.

D1. Nature of problems
P1. Focus
D2. Ego strength
P2. Modulation of affects
P3. Flexibility of interaction
P4. Self-concept in relation to ego ideal
D3. Self-observing capacity
P5. Response to trial interpretation
P6. Reflective ability
P7. Motivation
D4. Intelligence
P8. Intelligence
D5. Intrapsychic and interpersonal behavior
P9. Coping styles and cognitive skills
P10. Defense styles
P11. Personality traits
P12. Interpersonal relationships
D6. Overall suitability[1]
P13. Overall suitability[1]

D1-D6 = Suitability domains, P1-P12 = Suitability predictors
[1] Suitability measures covering more than one aspect of the individual suitability measures P1-P12.

reasons (de Jonghe et al. 2012). Furthermore, as patient selection via the consideration of indications and contraindications for a patient's analyzability − i.e., the capacity and motivation to form, maintain, and eventually relinquish a therapeutic relationship − has always been considered a crucial part of successful PA, randomization of potentially unsuitable patients for PA is generally not accepted or considered feasible by advocates of PA (American Psychiatric Association 1985, Caligor et al. 2012). Recently, findings supporting the feasibility of an RCT including PA and another therapeutic treatment, from the perspective of patient acceptance of randomization for PA, have, however, been presented (Caligor et al. 2012).

The only published RCT to date to include PA is the Munich Psychotherapy Study (MPS), which compared the effectiveness of PA (N = 43) to long-term psychodynamic psychotherapy (LTPP; N = 35) in the treatment of patients suffering from depressive disorders, and showed PA to be more effective than LTPP at a 3-year post-treatment follow-up with respect to depressive and global psychiatric

symptoms, personality functioning, and social relations (Huber et al. 2013). A few years later, a long-term cognitive-behavioral therapy (LTCBT) group was added to the MPS and patients referred to LTCBT were compared to patients randomized for PA or LTPP in a quasi-experimental design, which also demonstrated the superiority of PA in comparison to LTCBT (Huber et al. 2012). However, the average dose of PA (234 1-hour sessions over 39 months), as well as LTPP (88 1-hour sessions over 34 months), in the MPS was relatively low (de Maat et al. 2013), and is not considered psychoanalysis proper by most psychoanalysts or by the American Psychoanalytic Association, which sets the minimum frequency of PA to four sessions a week (de Jonghe et al. 2012). A quasi-experimental trial carried out by the Helsinki Psychotherapy Study (HPS) compared the effectiveness of a PA proper (an average of 646 45-50 minute sessions during 56 months, N = 41), a higher dose LTPP (an average of 232 45-50 minute sessions during 31 months, N =128), and two short-term therapies, short-term psychodynamic psychotherapy (STPP; an average of 19 45-50 minute sessions during 6 months, N =101) and solution-focused therapy (SFT; an average of 10 1.5-hour sessions during 7.5 months, N = 97), among depressive and anxiety disorder patients selected for PA based on their suitability and randomized for LTPP, STPP, and SFT, adjusting for potential confounding by the suitability variables evaluated in all patients (Knekt et al. 2011a). Similarly to the MPS, the HPS found PA more effective than LTPP with respect to depressive symptoms at the end of the PA in a 5-year follow-up from the start of the treatments. The HPS, however, found LTPP more effective than PA at the end of the LTPP in a 3-year follow-up from the start of the treatments, and short-term therapies more effective than PA after they had ended in a 1-year follow-up, suggesting that each therapy is at its best at the end of the treatment. Both the length of the treatment and the length of the follow-up thus seem important to the evaluation of the effectiveness of psychotherapies. The rest of the evidence on the effectiveness of PA is based on observational studies following a cohort of patients for whom PA is indicated, with or without a comparison group, and evaluating the pre-post treatment changes in various outcome measures. Optimally, the evidence on effectiveness is pooled across studies in a meta-analysis or summarized in a systematic review. Several reviews (Bachrach et al. 1991, Galatzer-Levy et al. 2000, Doidge 2001, Fonagy 2002, de Maat et al. 2009) but only one meta-analysis (de Maat et al. 2013) has been carried out on the effectiveness of PA. Although the study populations in the PA studies are small, typically under 50 patients, and consist mainly of patients suffering from severe mental disorders, patients undergoing PA have been shown to experience a significant symptom reduction and personality change upon treatment termination which have been found to remain stable in a post-treatment follow-up of up to 3.5 years (de Maat et al. 2013). Similarly to the one existing RCT (Huber et al. 2013) and quasi-experimental trial (Knekt et al. 2011a), the cohort studies comparing PA and long-term psychotherapies also found PA to be more effective in reducing psychiatric

symptoms (Kernberg et al. 1972, Kordy et al. 1988, Sandell et al. 2000, Berghout et al. 2012).

Similarly to PA, long-term psychotherapies have also been criticized for the lack of strong evidence-based support for their effectiveness. The effectiveness of long-term psychotherapies, LTPP in particular, has, however, been studied more extensively than that of PA, and several meta-analyses on the effectiveness of LTPP have appeared (Leichsenring and Rabung 2008, de Maat et al. 2009, Leichsenring and Rabung 2011, Smit et al. 2012, Leichsenring et al. 2013). Long-term psychotherapies are generally required to comprise at least 40-50 sessions, typically once or twice a week, and continue for at least one year, although they often last up to three years (Leichsenring and Rabung 2008, de Maat et al. 2009, Smit et al. 2012). The meta-analyses based on RCTs – comparing LTPP to other, typically less intensive, psychotherapeutic treatments or to non-psychotherapeutic treatment in routine clinical care, often referred to as treatment as usual – or cohort studies have generally supported the effectiveness of LTPP in the treatment of moderate to severe pathologies, and found it superior to shorter non-psychotherapeutic or psychotherapeutic treatments (Leichsenring and Rabung 2008, de Maat et al. 2009, Leichsenring and Rabung 2011). Furthermore, patients have been found to maintain the symptom reduction and personality change experienced during therapy, or even to continue to improve, in the post-treatment follow-up (Shedler 2010). One of the meta-analyses (Smit et al. 2012) focusing on RCTs and recovery rates, however, found LTPP superior only to non-psychotherapeutic, but not to psychotherapeutic, control treatments; the majority of the psychotherapeutic control treatments were, however, other forms of long-term psychotherapy, thus essentially showing LTPP to be as effective as other long-term psychotherapies (Leichsenring et al. 2013).

Studies on the effectiveness of long-term psychotherapies, as well as of PA, have mainly focused on the treatment of patients with complex mental disorders, such as chronic Axis I disorders, Axis II personality disorders, or multiple disorders, generally considered to be more treatable with long-term than with short-term psychotherapies (American Psychiatric Association 1994). Common mental disorders, particularly consisting of depressive and anxiety disorders, on the other hand, are considered treatable with both long-term and short-term psychotherapies. Long-term psychotherapies, which in many countries, including Finland, remain the most common treatment for such disorders (Konsensuskokous 2006, Grande et al. 2006, Beutel et al. 2012), have been shown to result in significant pre-treatment to post-treatment and post-follow-up changes (de Maat et al. 2009), but only a few studies have compared the effectiveness of different long-term psychotherapies or long-term and short-term psychotherapies in the treatment of depressive and anxiety disorders. The MPS compared the effectiveness of LTPP and LTCBT in the treatment of unipolar depression in a quasi-experimental design and found LTPP to be more effective than LTCBT in reducing interpersonal problems, but not psychiatric symptoms, both at the treatment termination and at the 3-year post-

treatment follow-up. Recently, the first RCT comparing LTPP and LTCBT has been initiated, and the RCT design will also be compared to a design in which the treatments are assigned according to patient preference (Beutel et al. 2012). In the HPS − the only RCT comparing the effectiveness of individual long-term, psychodynamic, and short-term, psychodynamic and solution-focused, therapies in the treatment of depressive and anxiety disorders − short-term therapies were found, on average, to reduce symptoms and improve work ability and self-concept faster, whereas long-term psychotherapy was found more effective in the long-term (Knekt et al. 2008a, b, Lindfors et al. 2012, Knekt et al. 2013).

Although short-term psychotherapies were the last to develop, the vast majority of the psychotherapy effectiveness studies have focused on them. Naturally, short-term psychotherapies, defined to last a maximum of 40 sessions but typically lasting 10-20 sessions (Cuijpers et al. 2013), are shorter in duration and thus easier to study. The number of effectiveness studies particularly increased following the development of short-term psychotherapies other than psychodynamic psychotherapies (see Section 2.1), which in essence claimed to challenge the effectiveness of short-term psychodynamic psychotherapies. A huge number of individual studies, systematic literature reviews, and meta-analyses comparing the effectiveness of different forms of short-term psychotherapies have appeared from the 1960s to date but, overall, the findings have supported 1) the effectiveness of different forms of individual short-term psychotherapies over a) being on a waiting-list for therapy (Gingerich and Eisengart 2000, Leichsenring and Leibing 2007, Abbass et al. 2011, Nieuwsma et al. 2012) and b) treatment as usual (Gingerich and Eisengart 2000, Churchill et al. 2001, Leichsenring and Leibing 2007, Wampold et al. 2011), and 2) equal effectiveness of a) different forms of individual short-term psychotherapies (Wampold et al. 2002, Leichsenring and Leibing 2007, Abbass et al. 2011, Braun et al. 2013, Cuijpers et al. 2013) and b) individual short-term psychotherapies and psychiatric medication (de Maat et al. 2008, Roshanaei-Moghaddam et al. 2011, Cuijpers et al. 2013). The accumulating research evidence was already summarized in 1975 by Luborsky et al. using the famous Dodo bird verdict "Everyone has won and all must have prizes" from Alice in Wonderland. Nonetheless, the debate over the equal effectiveness of different individual short-term psychotherapies has continued to the present day, despite the emergence of further research evidence in support of such equivalence.

Broadly, there are two opposing perspectives in this debate (Budd and Hughes 2009). Supporters of the first perspective have accepted the equal effectiveness of individual short-term psychotherapies and have concluded it to be due to the common factors present in any therapy, e.g., expectation of improvement, therapeutic alliance between patient and therapist, and providing a meaningful explanation for the patient's difficulties, having a bigger impact on the therapy outcome than any therapy-specific techniques or ingredients (Wampold 2001).

Supporters of the second perspective, on the other hand, consider common factors to be necessary but not sufficient elements of therapy process, and argue that certain therapies are more effective than others in the treatment of a specific disorder (Chambless and Hollon 1998); these effects are, however, masked in individual studies and meta-analyses analyzing the effectiveness of short-term psychotherapies in the treatment of various disorders. This perspective is particularly maintained by the proponents of CBT, which has a greater emphasis on causal models of the etiology of the disorder than other therapy orientations (Budd and Hughes 2009). This has directed effectiveness research more towards the study of one specific disorder at a time. No disorder-specific differences between CBT and other psychotherapies in the treatment of depressive and anxiety disorders have, however, been found in recent meta-analyses (Baardseth et al. 2013, Cuijpers et al. 2013).

A disorder-specific approach to effectiveness research, restricting the selection criteria of a study based on a diagnosis, may maximize homogeneity but is also likely to minimize the generalizability of findings to real-world clinical populations (Budd and Hughes 2009). In routine clinical practice, patients with the same primary diagnosis may present with different comorbidities and psychiatric histories. Definitions of different diagnoses also consist of partly overlapping symptoms and are thus not conceptually distinct entities. Furthermore, the nature and severity of a disorder is only one way of categorizing patients. Although a treatment may be shown to be effective in relation to a specific disorder on average, it cannot be assumed that all patients suffering from the disorder will benefit from the same treatment. In addition to patients' disorders, their psychological, intrapsychic and interpersonal, capacities, have also been considered major determinants of psychotherapy suitability and accordingly of the therapeutic outcome (see Section 2.2). Careful pre-selection of suitable patients, both based on their disorders and psychological capacities, has been emphasized since the development of psychoanalysis. Respectively, when the development and study of the first well-defined individual short-term psychodynamic psychotherapies began in the 1950s-1970s (see Section 2.1), distinguishing the patients suitable for them from those requiring longer treatment was considered requisite for a successful therapy outcome (see Section 2.2). At the time, it was thus well recognized, within the psychodynamic psychotherapy practice (Malan 1963), that the evaluation of patients' suitability for psychotherapy, including strengths and weaknesses in their psychological capacities, should precede the evaluation of the effectiveness of psychotherapy. If unsuitable patients are chosen for a particular psychotherapy there is a risk that no change or even a change for the worse will occur, in which case the patients' recovery will have been delayed and the resources have been wasted. The same decades, however, also witnessed the rise of several other forms of psychotherapies (see Section 2.1) whose advocates did not emphasize suitability for psychotherapy as much as the advocates of psychodynamic psychotherapies (see Section 2.2). Thus, along with the development of different psychotherapy forms,

the focus in the literature shifted from the study of psychotherapy suitability towards the study and comparison of the effectiveness of different psychotherapies. However, if the role and effect of suitability factors are not evaluated and taken into account, they are likely to confound the results of effectiveness studies, as the patients being compared may not be equally suitable for the treatments they receive, and accordingly the differences in their effectiveness may not be due to the therapy itself but rather due to differential patient selection. The effectiveness of different treatments should only be compared among patients considered suitable for them, and thus relevant suitability criteria should be included in the selection criteria of effectiveness studies.

Psychological psychotherapy suitability selection criteria have been generated by gradually refining the initial criteria, mainly based on clinical experience and preconceived ideas, as empirical patient data have accumulated (see Section 2.2). These data, and the evidence based on them, mainly consist of a collection of convincing case studies. More systematic research is thus needed to be able to arrive at firmer conclusions on the role of these suitability criteria. First, a potential criterion for suitability for a particular psychotherapy needs to be reliably measured, and thus a validated suitability measure – assessing what is intended and shown to be replicable by different, equally trained assessors over time – must be established for its evaluation. Second, for such a measure to be truly considered a suitability measure it also needs to be shown to predict the outcome of psychotherapy; patients with better values with respect to such a suitability measure should benefit more from psychotherapy than patients with poorer values. Furthermore, a suitability measure truly separating suitability for, and thus the outcome of, two different psychotherapies, such as short-term and long-term psychodynamic psychotherapies as originally intended, needs to predict a differential outcome in those two therapies, i.e., the patient needs to benefit more from one of the two therapies. If the patient characteristics that predict, and particularly separate, the outcome of psychotherapies can be reliably identified and integrated into the treatment referral process, therapeutic interventions can be made more rational, faster, and effective (Korchin and Schuldberg 1981, Maruish 1994). The evaluation and utilization of patients' suitability for different treatments thus has the potential to sharpen their outcome prognosis.

In the following, literature reviews on a) the validated psychological psychotherapy suitability assessment scales (Section 2.4), and b) the psychological psychotherapy suitability selection criteria as predictors of the outcome of individual short-term and long-term psychotherapies in the treatment of Axis I disorders (Section 2.5) are presented.

2.4 Review of psychotherapy suitability assessment scales

Patient suitability for psychotherapy can be evaluated based on an interview or patient self-report (Davanloo 1978, Fisher et al. 1999). Self-report suitability instruments are by and large faster and easier to administer. Self-evaluation of the suggested psychological psychotherapy suitability criteria (see Table 1) by a patient seeking therapy due to a psychiatric disorder may, however, be complicated in that the disorder may partly be related to the patient's aptitudes forming the suitability criteria, thus making the self-report sensitive to the influence of the actual psychopathology. The criteria suggested for indicating psychotherapy suitability also involve psychological processes that the individuals are not always aware of and the self-evaluation therefore only reflects their conscious derivates (Van et al. 2009). Moreover, many of the qualities relevant to psychotherapy suitability are considered as being manifested in the patient's way of relating to and interacting with others and are thus best determined via an interview. Some of the essential suitability criteria, i.e., the patient's flexibility of interaction and response to trial interpretation, can only be evaluated based on interaction between the patient and the interviewer. For these reasons, an interview-based suitability assessment carried out by a clinical observer, with knowledge of the content of the suitability concepts and training in their evaluation, is generally considered the most reliable and objective, and thus recommendable, method for the evaluation of complex psychotherapy suitability phenomena.

Several observer-rated and self-report instruments for the evaluation of different aspects of patient suitability have been developed, validated, and widely applied. However, most of the available instruments only address one aspect of suitability. A comprehensive suitability evaluation, covering all or most of the suggested suitability criteria based on such instruments, is likely to require a large amount of time, and may consequently limit findings and predictions to the most motivated and cooperative patients only (Fisher et al. 1999). A validated psychotherapy suitability assessment scale covering several aspects of psychological suitability within a relatively short timeframe would thus probably comprise an optimal method for the evaluation of a patient's psychotherapy suitability.

Here, a literature review on published interview-based psychotherapy suitability assessment scales is presented.

2.4.1 Selection of studies

The review covers published original studies on interview-based psychotherapy suitability assessment scales aimed at assessing the psychotherapy suitability of patients mainly suffering from Axis I disorders, chiefly based on an evaluation of psychological suitability (see Table 1), and covering more than one aspect of psychological suitability, for which reliability, validity, or prediction has been assessed and reported (Table 2). Studies were identified by checking the PubMed

Patient suitability for short-term and long-term psychotherapy

Table 2. Criteria for inclusion of studies for literature review on psychotherapy suitability assessment scales.

1. Scale is aimed to assess suitability for psychotherapy of patients mainly suffering from Axis I disorders.

2. Scale mainly assesses psychological suitability[1] and covers more than one aspect of psychological suitability.

3. Interview procedure is used to evaluate the scale.

4. Reliability, validity, or prediction of the scale has been assessed and reported.

[1] See Table 1.

databases and examining the reference lists of identified articles. A total of 28 studies on reliability, validity, or prediction of 11 separate interview-based psychotherapy suitability assessment scales fulfilling the selection criteria were identified; for one scale three different versions (Heiberg 1975; Husby et al. 1985, Husby 1985a, Barth et al. 1988a, b; Høglend et al. 1992a) and for another scale two different versions (Safran et al. 1993, Myhr et al. 2007) were presented (Tables 3 and 4).

2.4.1.1 Data extraction

The following data were extracted from each study: the first author's name, year of publication, name of the scale, form and length of therapy for which the suitability was evaluated, information sources used in the construction of the scale, the number of items on the scale, their rating scale, which of the 12 suitability aspects (see Table 1: P1-P12) they measured, whether a score summarizing overall suitability (Table 1: P13) was formed, and how many of the scale items evaluated patient characteristics other than psychological suitability (Table 3). Validation of the scales was evaluated and compared based on training for the suitability interview, the number of interview sessions during which the suitability was evaluated, the number of interviewers used, the sample size based on which the reliability of the scale was evaluated, reliability evaluation method/s, reliability results, the sample size based on which the validity of the scale was evaluated, validity evaluation method/s, validity results, sample size based on which the prediction by the scale was evaluated, prediction evaluation method/s, the form and length of therapy of the outcome of which the prediction by the scale was evaluated, the mean number and range of therapy sessions attended by the patients, and therapist training (Table 4). Information not available was marked as not reported (NR).

2.4.1.2 Standardization

As the studies included in the literature review reported variables of interest using different terms and different levels of detail, and applied different assessment

Table 3. Description of the psychotherapy suitability assessment scales.[1]

| Author | Scale | Therapy[3] | Influence | Items[4] | Number of suitability measures (P1-P13) from different domains (D1-D6) included in the scale[2] | | | | | | | | | | | | | |
					D1 P1	D2 P2	P3	P4	D3 P5	P6	P7	D4 P8	D5 P9	P10	P11	P12	D6 P13[5]	Other[6]
1A. Heiberg 1975	Initial Evaluation Form for Short-Term Anxiety-Provoking Psychotherapy (STAPP)	STPP(I)	Sifneos	5	1[7]	-	1	-	-	-	1[8]	1	-	-	-	1	1	-
1B. Husby et al. 1985, Husby 1985a, Barth et al. 1988a,b	Modified version 1 of selection criteria for STAPP			6	2[7]	-	1	-	-	-	1[8]	1	-	-	-	1	1	-
1C. Høglend et al.1992a	Modified version 2 of selection criteria for STAPP			14	2[7]	-	2	-	-	-	7[8]	1	-	-	-	1	-	1
2. Brodaty et al. 1982	Selection criteria	STPP(I)	Malan, Sifneos, Garfield & Affleck, Stoler, Wolberg, Small	6	1	1	-	-	1	1	1	-	-	-	1	-	1	-
3. Persson & Alström 1983	A scale for rating suitability for insight-oriented psychotherapy	STPP(I)	Dewald	24	1	6	1	-	-	1	3	1	2	1	2	1	-	5
4. Buckley 1984	Rating scales	STPP(I)	NR	59	-	5	-	-	-	-	-	-	8	16	16	5	-	9
5. Piper et al. 1985	-	STPP(I)	NR	15	2	2	1	-	1	1	2	-	-	1	1	4	-	-
6. Vaslamatzis & Verveniotis 1985	Selection form	STPP(I)	Davanloo	5	2	-	-	-	1	-	1	-	-	-	-	1	1	-
7. Alpher et al. 1990	Capacity for Dynamic Process Scale (CDPS)	STPP(I)	Dewald	9	-	3	1	-	-	3	-	-	-	-	1	1	1	-
8A. Safran et al. 1993	Suitability for Short-Term Cognitive Therapy (SSCT) Rating Scale	STCT	Malan, Mann, Strupp & Binder, Sifneos, Davanloo, Bordin	9[9]	1	1	1	-	-	1	2[9]	-	-	1	-	1	1	1
8B. Myhr et al. 2007	Modified version of SSCT			10[9]	1	1	1	-	-	1	3[9]	-	-	1	-	1	1	1

No. & Author	Scale	Therapy[3]	Ref.[1]	Items[4]															
9. Rosenbaum et al. 1997	Dynamic Assessment Interview (DAI)	PP	Selzer	10	-	2	-	-	1	2	2	-	-	-	-	-	2	2	-
10. Fisher et al. 1999	Systematic Treatment Selection (STS) Clinician Rating Form	SEVERAL	NR	226[10]	-	-	-	-	-	-	-	1	-	1	-	1	-	-	1
11. Sigal et al. 1999	Checklist for Brief Psychodynamic Therapy	STPP(I)	NR	39[11]	1	3	-	1	-	1	8	-	-	-	-	-	5	-	13

NR = Not reported

[1] For selection criteria see Table 2.

[2] *D1. Nature of problems:* P1. Focus
D2. Ego strength: P2. Modulation of affects, P3. Flexibility of interaction, P4. Self-concept in relation to ego ideal
D3. Self-observing capacity: P5.Response to trial interpretation, P6. Reflective ability, P7. Motivation
D4: Intelligence: P8. Intelligence
D5. Intrapsychic and interpersonal behavior: P9. Coping styles and cognitive skills, P10. Defense styles, P11. Personality traits, P12. Interpersonal relationships
D6. Overall suitability: P13. Overall suitability
More detailed description of the individual suitability measures is given in Appendix 1.

[3] PP = Psychodynamic Psychotherapy, STCT = Short-Term Cognitive Therapy, STPP(I) = Short-Term Psychodynamic Psychotherapy (Interpretative).

[4] Number of individual items included in the scale: psychological (P1-P12)[2] or other[6].

[5] Overall assessment of suitability (e.g., sum or mean of the individual suitability items)[2].

[6] Other than psychological suitability measures (e.g., symptoms).

[7] Two aspects are considered in the evaluation of focus: 1) patient's ability to give priority to one out of a variety of complaints, and 2) evaluator's ability to recognize that what underlies this complaint is associated with a circumscribed field of psychological conflicts that are triangular, i.e., stemming from the oedipal period, rather than dyadic, i.e., stemming from pre-oedipal period. Later modifications of this scale (Husby et al. 1985, Husby1985a, Barth et al. 1988a, b, and Høglend et al.1992a) separated these two aspects into two separate questions.

[8] Seven aspects are considered in the evaluation of patient's motivation: 1) ability to recognize that symptoms are psychological, 2) introspection and honesty in reporting about self, 3) willingness to participate actively in treatment, 4) active curiosity about self, 5) desire to change, 6) expectations of results realistic, 7) willingness to make reasonable sacrifices. A later modification of this scale by Høglend et al. (1992a) separated these seven aspects into seven separate questions.

[9] A later modification of this scale (Myhr et al. 2007) included one more suitability item reflecting patient's motivation.

[10] 226 items were grouped into three factors, individuals items were not described.

[11] Only 32 of the 39 items were reported.

Table 4. Reliability, validity, and prediction of the psychotherapy suitability assessment scales.[1]

Author	Items[2]	Rating	Sessions[4]	Training[5]	Reliability			Validity		Prediction		
					N_1[6]	N_2[7]	Inter-rater[8]	N_1[9]	Type[10]	N_1[11]	Therapy[12]	Type[13]
1A. Heiberg 1975, 1976	5	2-point	S: 1	S: T	167	2	0.50-0.98[14]	167	C, D	-	-	-
1B. Husby et al. 1985, Husby 1985a	6	3-point	S: 1 T: 20 (4-39)	S: T T: T	-	-	-	-	-	36, 33	STPP(I): STAPP, BDP	O
Barth et al. 1988a, b			S: 1 T: STAPP ≤ 20 BDP ≤ 40 FIAT 25-30	S: T T: T	44	4	0.63-1.00[15]	-	-	44, 34	STPP: STAPP, BDP, FIAT	C, O
1C. Høglend et al. 1992a, Høglend et al. 1992b, Høglend 1993, Høglend et al. 1993, 1994	14	8-point	S: 1 T: 28 (9-53) S: 1 T: 27 (9-53) S: 1 T: 28 (9-53)	S: T T: T S: T T: T S: T T: T	10	4	ICC: 0.25-0.52	-	-	34, 39 42, 45 40, 43	STPP(I) STPP(I) STPP(I)	O O O
2. Brodaty et al. 1982	6	9-point	S: 1-2 T: 8	S: NR T: T	-	-	-	-	-	18	STPP(I)	O
3. Persson & Alström 1983	24	5-point	S: 1 T: 12-24	S: NR T: NR	25	2	CC: 0.08-0.85	-	-	38	STPP(I), STPP(S)	C
4. Buckley et al. 1984	59	4-point	S: NR T: 43	S: NR T: T	NR	2	K: 0.89	-	-	21	STPP(I)	O
5. Piper et al. 1985	15	7-point	S: 2 T: 22	S: NR T: T	15	2	CC: 0.14 (0.02-0.45)	-	-	21	STPP(I): BDP	O
de Carufel & Piper 1988			S: 1-2 T: NR	S: NR T: T	-	-	-	-	-	79	STPP: IND, GRP LTPP: IND, GRP	O
6. Vaslamatzis & Verveniotis 1985, Vaslamatzis et al. 1989	5	5-point	S: NR T: ≤ 30 S: 3-5 T: ≤ 30	S: NR T: ST S: NR T: NR	-	-	-	-	-	24 27	STPP(I) STPP(I)	C C

Study	Items[2]	Scale	Training (S/T)[5]	Sessions (S/T)[3,4]	n[6]	Raters[7]	Reliability[8]	n[9]	Validity[10]	n[11]	Therapy[12]	Outcome[13]
7. Alpher et al. 1990	9	9-point	S: NR, T: T	S: 1, T: 23	25	2	ICC: 0.78[16]	-	-	25	STPP(I)	O
Baumann et al. 2001,		5-point	S: T, T: ST	S: 1, T: 23	38	2	ICC: 0.64-0.89	38	D	35	STPP(I)	C
Cromer & Hilsenroth 2010		5-point	S: T, T: ST	S: 1, T: 27	71	2	ICC: 0.50-0.72	-	-	71	STPP(I)	C, O
8A. Safran et al. 1993	9	9-point	S: T, T: NR	S: 1, T: ≤ 20	11	3	ICC: 0.78 (0.46-0.98)	NR	C, D	32-42	STCT	O
8B. Myhr et al. 2007	10	5-point	S: NR, T: T, ST	S: 1, T: 19 (3-65)	28	7	ICC: 0.70 (0.60-0.92)	-	-	113	STCT	O
9. Rosenbaum et al. 1997,	10	6-point	S: T	S: 2-3	67	2	ICC: 0.49-0.88	-	-	-	-	-
Jørgensen et al. 2000,		6-point	S: T, T: T, ST	S: 1, T: 69 (20-135)	-	-	ICC: 0.88[16]	20	C	20	LTPP	O
Valbak et al. 2004		5-point	S: T, T: T, ST	S: 2-3, T: 78 (32-289)[17]	-	-	-	-	-	74[17]	PP[17]	O
10. Fisher et al. 1999	226	2-point	S: T, T: T, ST	S: 1, T: NR	46	5	K: 0.84 (0.77-0.99)	89-204	C, D	289	NR[18]	(O)[18]
11. Sigal et al. 1999	39	2-point	S: NR, T: T, ST	S: 1, T: 12 (12-12)	-	-	-	-	-	43	STPP(I)	O

NR = Not reported
1 For selection criteria see Table 2.
2 Number of individual suitability items included in the scale (see Table 3).
3 Number of interview sessions used for evaluating the suitability scale; Therapy (T): The mean number (range) of therapy sessions provided.
4 Suitability (S): The number of interview sessions used for evaluating the suitability scale. Therapy (T): The mean number (range) of therapy sessions provided.
5 Suitability (S): Training for the evaluation of the scale: T=Trained, NT= Untrained; Therapist (T): Training for the therapy form: T=Trained, ST=Semi-trained, NT= Untrained.
6 Number of patients assessed for their suitability in the reliability study.
7 Number of raters who assessed the suitability of each patient.
8 Agreement between individual raters' assessments: CC=Correlation coefficient, ICC=Intraclass correlation coefficient, K=Kappa coefficient. Mean (Range) unless otherwise mentioned.
9 Number of patients included in the validity study.
10 Criterion (C) or discriminating (D) validity. Underlined when validity supported.
11 Number of patients evaluated in the prediction study, at the end of therapy and/or at follow-up. Barth: end of therapy and 2-year follow-up, Husby: 2- and 5-year follow-up, Høglend: 2- and 4-year follow-up.
12 The type of therapy provided in the study: BDP = Brief Dynamic Psychotherapy (Malan), FIAT = an integrative form of short-term psychodynamic psychotherapy applying supportive, behavioral, and cognitive procedures (Nielsen), GRP = Group, IND = Individual, LTPP = Long-Term Psychodynamic Psychotherapy, PP = Psychodynamic Psychotherapy, STAPP = Short-Term Anxiety-Provoking Psychotherapy (Sifneos), STCT = Short-Term Cognitive Therapy, STPP = Short-Term Psychodynamic Psychotherapy, STPP(I) = STPP (Interpretative), STPP(S) = STPP (Supportive).
13 Compliance (C) in the therapy form assessed suitable or prediction by the suitability scale of therapy outcome (O). Underlined when compliance / outcome prediction supported.
14 95% confidence interval for range of accordance.
15 Statistical method not explicitly stated.
16 For overall suitability score, i.e., sum of the values of the individual suitability items.
17 61 long-term therapies, 13 short-term therapies (12-24 sessions).
18 Reference to unpublished manuscript by Beutler et al. which remains unpublished.

methods, the following standardization and grouping was carried out. A variety of names and abbreviations for the same therapy form, psychodynamic psychotherapy in particular, were used in the original articles, and the therapy names and their abbreviations were therefore standardized. The psychological suitability variables included in the scales were grouped under the 12 previously presented suitability selection criteria, which were further placed under 5 suitability domains, according to their clinically relevant conceptual scope (Table 1, Appendix 1). Different approaches to the evaluation of overall suitability were grouped under the overall suitability domain and variable (Table 1, Appendix 1). The training of the interviewers was categorized into two groups: trained and untrained. The training of the therapists was categorized into three groups: therapists adequately trained and experienced in the therapy they provided in the study, therapists under training or supervision for the therapy they provided in the study, and untrained therapists with potential training and/or experience in other therapy forms but not the one they provided in the study.

2.4.2 Overview of the selected studies

The first of the 11 identified psychotherapy suitability assessment scales was published in 1975 (Heiberg 1975) and the rest of the scales were introduced in the 1980s and 1990s (Table 3). The two sources of information generally used in the construction of these scales were 1) clinical experience and beliefs of factors relevant for psychotherapy suitability, and accordingly psychotherapy outcome, as reflected in clinical writings, and 2) research evidence on factors predictive of psychotherapy outcome. The most influential clinical writings included the work by Sifneos (1967, 1968a, b, 1972, 1979), Malan (1963, 1976a, b), Davanloo (1978, 1980), and Dewald (1964) (see Section 2.2, Table 3). Other writings that guided the construction of the scales included the work by Garfield and Affleck (1961), Stoler (1963), Wolberg (1965, 1977), Small (1971), Mann (1973), Bordin (1979), Strupp and Binder (1984), and Selzer et al. (1987). Most of these authors were dynamic therapists interested in identifying patients suitable for interpretative short-term psychodynamic psychotherapy (STPP). Accordingly, most of the proposed suitability assessment scales were either intended for the evaluation of suitability for STPP or influenced by the suitability criteria first developed within psychodynamic psychotherapy (Table 3). One scale was developed for the evaluation of suitability for both short-term and long-term psychodynamic psychotherapy (Rosenbaum et al. 1997, Jørgensen et al. 2000, Valbak et al. 2004), one scale for the evaluation of suitability for short-term cognitive therapy (STCT; Safran et al. 1993, Myhr et al. 2007) and one scale for the evaluation of suitability for different kinds of treatments (Fisher et al. 1999). The last scale was the most extensive one consisting of 226 individual items. The majority (7/11) of the other scales were quite concise, consisting of 5-15 items, whereas the rest of the scales consisted of 24-59 items

(Table 3). Half of the scales, however, also included items assessing aspects of the patients other than their psychological suitability (Table 3), such as symptoms, which makes it difficult to evaluate the independent role of psychological suitability when such scales are used to predict psychotherapy outcomes. All of the longer scales included several such items, whereas the more concise scales only included one non-suitability item. The most commonly evaluated aspects of psychological suitability were modulation of affects, motivation, and interpersonal relationships (all in 8/11 scales), and the focality of the problems, reflective ability, and personality traits, although these were evaluated from very different aspects (7/11 scales). Flexibility of interaction (5/11), response to trial interpretation and defense styles (4/11), coping styles (3/11), intelligence (2/11), and self-concept (1/11) were evaluated for a minority of the scales. Where several versions of the same scale existed (Heiberg 1975, Husby et al. 1985, Husby 1985a, Barth et al. 1988a, b, Høglend et al. 1992a; Safran et al. 1993, Myhr et al. 2007), the modified versions either divided an item from the original scale into two or more separate items, to better address the different aspects of that suitability dimension (Husby et al. 1985, Husby 1985a, Barth et al. 1988a, b, Høglend et al. 1992a) or added new items to the original scale (Myhr et al. 2007, Høglend et al. 1992a) (Table 3). Later modifications of the first published suitability assessment scale (Heiberg 1975), based on Sifneos' (1972) selection criteria for Short-Term Anxiety-Provoking Psychotherapy (STAPP), also shifted from the evaluation of intelligence to the evaluation of problem-solving capacity (Barth et al. 1998a, b), to avoid an overly cognitive or IQ-based conception of intelligence (Barth et al. 1988b); problem-solving capacity as defined by Høglend et al. (1992a, 1994) included evaluation of self-understanding and insight, thus closely relating to reflective ability (Table 3). In over half of the published studies on psychotherapy suitability assessment scales, the evaluation of individual aspects of suitability was followed by an overall suitability assessment or calculation of a suitability score, typically summing up the values of the individual suitability items; in one study (Rosenbaum et al. 1997) both an overall assessment of positive and negative suitability aspects and an overall sum of the values of the individual suitability items was provided (Table 3).

2.4.3 Reliability of the psychotherapy suitability assessment scales

Reliability is the degree of consistency and stability of results based on an assessment tool. Reliability of an interview-based assessment scale is typically evaluated by measuring inter-rater reliability or test-retest reliability. Inter-rater reliability or agreement measures how well two or more raters agree in their assessments. Test-retest reliability measures how repeatable an individual rater's assessments are over time. Of the 11 published interview-based psychotherapy suitability assessment scales, 8 have been studied for reliability (Table 4). Agreement between individual raters' assessments was studied for all 8 scales

(Heiberg 1976, Barth et al. 1988a, b, Høglend et al. 1992a; Persson and Alström 1983; Buckley et al. 1984; Piper et al. 1985; Alpher et al. 1990, Baumann et al. 2001; Safran et al. 1993, Myhr et al. 2007; Rosenbaum et al. 1997; Fisher et al. 1999) but the repeatability of individual raters' assessments over time was studied for none.

In most of the studies, suitability was assessed based on one, usually semi-structured, interview, lasting from 45 minutes to 2 hours (Table 4); only a few scales were assessed based on several interview sessions (Brodaty et al. 1982, Piper et al. 1985, Vaslamatzis et al. 1989, Rosenbaum et al. 1997) although reliable evaluation of some aspects of suitability has been considered to require more than one interview (see Section 2.2). In the reliability studies, the suitability assessment interview carried out by the primary assessor was typically videotaped and the videotaped interview was then evaluated by the secondary assessor/s. In two studies, all raters made their ratings from audiotapes (Safran et al. 1993) or videotapes (Rosenbaum et al. 1997); in the study by Rosenbaum et al. (1997) the ratings were made based on only the first assessment interview although 2-3 assessment interviews were carried out. Only rarely was the same patient interviewed by more than one assessor (Heiberg 1975). The number of assessors evaluating the suitability of each individual patient varied from 2 to 7 and the number of patients assessed in the reliability studies from 10 to 167, although one study (Buckley et al. 1984) failed to report the number of patients assessed (Table 4). The training performed for the assessment of the suitability scale was not reported for the majority of the scales (Table 4), and in the studies reporting the training its duration varied from 10 hours (Fisher et al. 1999) to 5 years (Safran et al. 1993). In many cases, weekly workshops, including evaluation of patients' suitability followed by discussion about suitability scoring and interpretation guidelines, were organized throughout the study period (Heiberg 1975: 6 months, Husby et al. 1985: 3 years; Baumann et al. 2001: 22 months), similarly to the original approach by Malan and Sifneos (see Section 2.2). Some studies mentioned special training for the assessment approach as part of postgraduate training (Barth et al. 1988a, b, Cromer and Hilsenroth 2010), training in the manual for the assessment interview (Rosenbaum et al. 1997, Jørgensen et al. 2000), or the practice of the approach with a certain number of patients (Barth et al. 1988a: 20 patients, Høglend et al. 1992a: 10 patients), with more training improving the reliability of the assessment. One study compared the ratings of interviewers specifically trained for the suitability assessment and clinical interviewers with relatively little training in the method, and concluded that these ratings did not correlate highly (Valbak et al. 2004: r = 0.31).

Three statistical methods were used in the evaluation of inter-rater reliability: the correlation coefficient (CC), intra-class correlation coefficient (ICC), and kappa coefficient. The CC is a measure of association between two assessments, with a CC of 1 representing a perfect positive linear relationship and a CC of 0 no linear relationship. Strong association can, however, exist without strong agreement, for

example when one rater's assessments are consistently two points higher than another rater's assessments, and thus the CC is not a good measure of agreement. The ICC and kappa coefficient, on the other hand, are measures of agreement. The ICC takes account of both the correlation between the assessments and differences in the assessments (Gwet 2012). The kappa coefficient further accounts for the amount of agreement between raters that is expected to occur through chance. Both unweighted (Cohen 1960) and weighted (Cohen 1968) forms of kappa have been defined. When quadratically weighted kappa coefficients are used, the results are equivalent to those obtained using the ICC (Krippendorf 1970, Brenner and Kliebsch 1996). Different ranges of arbitrary values for the ICC and kappa have been suggested in order to describe the strength of agreement they represent. There is, however, wide disagreement about the usefulness of such benchmarks, and even about the kappa statistic itself, since it varies depending on the number of categories and weighting scheme used; results based on different scales and different weighting schemes should thus be compared with caution. Bearing this in mind, in the following description and comparison of the results for pragmatic reasons inter-rater agreement is considered poor, fair, good, and excellent for values 0.00-0.39, 0.40-0.59, 0.60-0.74, 0.75-1.00, respectively (Fleiss 1981).

The agreement between individual raters' assessments was reported separately for the individual suitability items of 4 scales (Persson and Alström 1983; Safran et al. 1993; Rosenbaum et al. 1997; Baumann et al. 2001, Cromer and Hilsenroth 2010). The consistently highest agreement was found for the items derived from the Suitability for Short-Term Cognitive Therapy (SSCT) scale (Safran et al. 1993); of the 9 items, all except focality showed excellent agreement beyond chance between the three assessors. The sample size (N = 11) in this study was, however, among the smallest of all the reliability studies. A later study (Myhr et al. 2007) evaluating the reliability of the same scale in a larger patient sample (N = 28) demonstrated slightly lower average agreement (0.70 versus 0.78 in Safran et al. 1993) but the range of ICCs (0.60-0.92) indicated good agreement or excellent agreement between seven raters for all suitability items. For the Capacity for Dynamic Process Scale (CDPS; Baumann et al. 2001, N = 38), 6 of the 9 scale items showed excellent agreement and the remaining 3 items good agreement between two assessors. However, in a later study by Cromer and Hilsenroth (2010) on the reliability of the same scale in a larger patient sample (N = 71), the 6 items showing excellent agreement and the 3 items showing good agreement in the study by Baumann et al. (2001) showed good and fair agreement, respectively. Of the 9 individual items of the Dynamic Assessment Interview (DAI; Rosenbaum et al. 1997), psychological mindedness showed excellent agreement (ICC = 0.80), 6 items good agreement (ICC = 0.68-0.74), and 2 items (confidence in treatment, attractiveness) fair agreement between two assessors. The assessments in this study were based on 67 patients, which was among the largest sample sizes used in the reliability studies. Of the 24 items of the scale for rating suitability for insight-oriented psychotherapy (Persson and Alström

1983), self-confidence and possibility of psychodynamic formulation with a circumscribed focus had the highest, excellent or nearly excellent, agreement; all three motivation items (willingness to get rid of the symptoms, willingness to work actively with oneself, and ability to take a reasonable amount of responsibility), on the other hand, had poor agreement, and the worst agreement was identified for intelligence. Agreement was, however, measured based on Spearman rank correlations which evaluate the agreement in ranking, not the magnitude, of assessments. Finally, agreement on the overall suitability was found excellent for both of the two scales reporting it (Rosenbaum et al. 1997: ICC = 0.81 (overall assessment) and 0.88 (DAI score), Jørgensen et al. 2000: ICC = 0.88 (DAI score); Alpher et al. 1990: ICC = 0.78, Baumann et al. 2001: ICC = 0.89, Cromer and Hilsenroth 2010: ICC = 0.69 (CDPS score)).

For the remaining 4 scales, either the range of the 95% confidence intervals for accordance (Heiberg 1976), the range of Pearson correlations (Piper et al. 1985), the range of ICCs (Barth et al. 1988a, Høglend et al. 1992a), the range of kappa coefficients (Fisher et al. 1999), or the median kappa coefficient (Buckley et al. 1984) for different suitability items was reported; they showed, on average, excellent or good agreement with the exception of two studies (Piper et al. 1985, Høglend et al. 1992a) which showed poor agreement (Table 4). Høglend et al. (1992a), however, demonstrated that the use of ICCs averaged over four raters significantly improved the reliabilities (ICCs for single interchangeable raters 0.25-0.52 versus ICCs averaged over four raters 0.38-0.83). The average ratings demonstrated excellent agreement on the patient's quality of interpersonal relationships (ICC = 0.83) and realistic expectations (ICC = 0.75), good agreement for 8, and fair or nearly fair agreement for only 3 of the remaining 11 psychological suitability items. The least reliable items included two of the seven motivation items (willingness to make reasonable sacrifices, sincerity/openness) and problem-solving capacity (use of self-understanding).

2.4.4 Validity of the psychotherapy suitability assessment scales

Validity is the degree to which the assessment tool measures what it theoretically aims to measure. Validity of an assessment scale is typically evaluated by measuring the association between the scale and an appropriate criterion measure, tapping theoretically similar dimensions to the scale (i.e., criterion or convergent validity), or the association between the scale and a discriminating measure, assessing theoretically a different phenomenon (i.e., discriminating or divergent validity). Of the 8 published interview-based suitability assessment scales studied for reliability, 5 were also studied for their validity (Heiberg 1976, Safran et al. 1993, Fisher et al. 1999, Jørgensen et al. 2000, Baumann et al. 2001) (Table 4). Both criterion validity (Heiberg 1976, Safran et al. 1993, Fisher et al. 1999, Jørgensen et al. 2000) and

discriminating validity (Heiberg 1976, Safran et al. 1993, Fisher et al. 1999, Baumann et al. 2001) were measured in 4 of the 5 studies evaluating validity.

Jørgensen et al. (2000) evaluated the criterion validity by measuring correlations between the DAI ratings of the patients interacting with the therapist in the assessment interview and the Structural Analysis of Social Behavior (SASB) rating of the same interpersonal process, but found no relationship. Baumann et al. (2001) evaluated the discriminating validity comprehensively, by measuring the correlation between the CDPS and three DSM-IV Axis V therapist rating scales (the Global Assessment of Functioning, the Global assessment of Relational Functioning, and the Social and Occupational Functioning Assessment Scale) and three self-report measures of global psychopathology and impairment in social and interpersonal functioning (the Symptom Checklist-90 Revised, the Social Adjustment Scale, and the Inventory of Interpersonal Problems): the CDPS was divergent from all these measures, thus supporting the discriminating validity. Fisher et al. (1999) studied the criterion validity by comparing the associations between the four subscales of the Systematic Treatment Selection (STS) Clinician Rating Form (subjective distress, externality, internality, resistance) and the similar subscales of the Minnesota Multiphasic Personality Inventory, and found statistically significant correlations between all respective subscales. To evaluate the discriminating validity, instead of using an external discriminating measure, Fisher et al. (1999) studied the intercorrelations between the four STS subscales, one of which can be considered to relate more closely to symptoms (subjective distress) and the others to psychological suitability; however, the symptom-related subscale was correlated with the suitability-related subscales, thus not strongly supporting the discriminating validity. Similarly, Heiberg (1976) studied the criterion validity by evaluating the relationships between the five psychological suitability questions, which were strongly related to one another, and the discriminating validity by evaluating the relationships between the psychological suitability questions and two questions assessing whether or not the patient was in a state of acute stress or crisis, which were not strongly related, thus concluding that both the criterion and the discriminating validity of the scale were supported. Safran et al. (1993) evaluated the validity by examining the relationships between the nine items and the mean score of the SSCT and the self-report measure of therapeutic alliance (Working Alliance Inventory (WAI)): the only significant correlation with the WAI was found with respect to the in-session alliance potential item, measuring the flexibility of the interaction between the patient and the therapist, which was seen to support both the criterion and the discriminating validity. Traditionally, however, the criterion and discriminating validity are evaluated using measures assessed at the same time; as the WAI in the study by Safran et al. (1993) was administered at the end of the third therapy session, the evaluation was inclined towards prediction rather than validity. For example Baumann et al. (2001) evaluated the relationship between the CDPS, evaluated prior to the start of the therapy, and the patient- and therapist-rated

alliance, evaluated both after the suitability assessment and after the third treatment session (the interviewer was also the treating therapist), but the purpose of these analyses was to evaluate the ability of the CDPS to predict early changes in therapeutic alliance, which it did not appear to predict.

2.4.5 Prediction by the psychotherapy suitability assessment scales

The prediction by the psychotherapy suitability assessment scale is typically evaluated by assessing how big a proportion of the patients complete the therapy, evaluated as being suitable for them, or by comparing how much the well-being of more and less suitable patients improves during the treatment and post-treatment follow-up. Of the 11 published interview-based psychotherapy suitability assessment scales, 10 were studied for their prediction (i.e., predictive validity): 4 in relation to therapy completion within the assigned therapy mode (Persson and Alström 1983; Vaslamatzis and Verveniotis 1985, Vaslamatzis et al. 1989; Barth et al. 1988a; Baumann et al. 2001, Cromer and Hilsenroth 2010) and 8 in relation to the change in well-being during and after therapy (Husby et al. 1985, Husby 1985a, Barth et al. 1988a, Høglend et al. 1992a, b, 1993, 1994, Høglend 1993; Brodaty et al. 1982; Buckley et al. 1984; Piper et al. 1985, de Carufel and Piper 1988; Alpher et al. 1990, Cromer and Hilsenroth 2010; Safran et al. 1993, Myhr et al. 2007; Jørgensen et al. 2000, Valbak et al. 2004; Sigal et al. 1999) (Table 4).

The suitability rating used in assigning suitable treatment or predicting the psychotherapy outcome was either an individual rater's rating (Brodaty et al. 1982; Persson and Alström 1983; Buckley et al. 1984; Piper et al. 1985, de Carufel and Piper 1988; Safran et al. 1993; Sigal et al. 1999), a mean score of several raters' ratings (Høglend 1992a, 1993, 1994, Høglend 1993; Alpher et al. 1990, Cromer and Hilsenroth 2010; Jørgensen et al. 2000, Valbak et al. 2004; Myhr et al. 2007), or a group rating reached after individual ratings (Barth et al. 1988a, b). Baumann et al. (2001) carried out two ratings and used both of them in the analysis. Husby et al. (1985) and Vaslamatzis and Verveniotis (1985) also carried out two ratings but did not report whether an individual rating or a combined rating was used in the analysis. The reliability of averaged ratings has been shown to be higher than the reliability of individual ratings, and it has thus been suggested that averaged ratings are more valid (Horowitz et al. 1979); however, in real life, patients are rarely evaluated by more than one person. In three studies the suitability rating was carried out after 1-5 therapy sessions (Sigal et al. 1999: after the first therapy session; Brodaty et al. (1982): at the end of the second therapy session; Vaslamatzis and Verveniotis 1985, Vaslamatzis et al. 1989: after 3-5 therapy sessions), whereas in the rest of the studies the suitability rating was carried out during 1-3 pre-therapy suitability assessment sessions. Also in these studies, the rater of suitability and the treating therapist were the same person either for all of the patients (Piper et al. 1985, de Carufel and Piper 1988; Høglend 1992a, 1993, 1994, Høglend 1993;

Bauman et al. 2001, Cromer and Hilsenroth 2010) or some of the patients (Barth et al. 1988a, b; Safran et al. 1993; Jørgensen et al. 2000, Valbak et al. 2004). Husby et al. (1985) and Myhr et al. (2007) did not explicitly report whether the suitability rating was carried out by the treating therapist, but this was apparently the case. Persson and Alström (1983), Buckley et al. (1984), and Alpher et al. (1990) did not provide the information required to judge whether the suitability rating and psychotherapy treatment were carried out by the same person or not. When the suitability rater is the therapist, the correlations with the therapist-rated psychotherapy benefit have been shown to be stronger, thus suggesting a bias towards giving the patient a higher rating if the therapist has previously given the patient a higher suitability score (Valbak et al. 2004).

Persson and Alström (1983) estimated predictive validity by evaluating whether the treatment of 7 patients assessed as being suitable for short-term insight-oriented psychotherapy could be carried through as planned, and whether 31 patients assessed as being unsuitable for insight-oriented psychotherapy during the course of their short-term supportive psychotherapy gave the impression that they would have profited from a more insight-oriented approach. As almost half of the patients in insight-oriented therapy were moved onto supportive therapy and a third in supportive therapy were considered to have possibly benefited from insight-oriented therapy, the results on the prediction of the scale were unconvincing. Vaslamatzis and Verveniotis (1985), on the other, found that a high suitability score and good values in 3 out of 5 individual suitability items (i.e., early transference confrontation, circumscribed problem, and high motivation) were strongly associated with the continuation of brief dynamic psychotherapy. In a later study, Vaslamatzis et al. (1989) further demonstrated that patients with better suitability scores were more likely to complete the therapy and less likely to drop out or continue their treatment beyond the intended length. Barth et al. (1988a) studied 44 patients, 10 of whom were assigned to the most demanding Short-Term Anxiety-Provoking Psychotherapy (STAPP; Sifneos 1972, 1978), 22 to the less demanding Brief Dynamic Psychotherapy (BDP; Malan1963, 1976a, b), and the remaining 12 patients to the least demanding, more integrative form of brief psychotherapy, the so-called FIAT model (Nielsen et al. 1984), allowing supportive, behavioral, and cognitive procedures (see Section 2.2). Assigned therapy was completed by 78%; two patients were changed from STAPP to BDP and four were considered to be in need of long-term therapy, whereas other patients ended their therapy prematurely due to other reasons, thus supporting the prediction of this assessment form. Baumann et al. (2001) found that the CDPS score effectively differentiated between those patients who terminated and those who continued short-term psychodynamic psychotherapy: patients with better score values were more likely to complete the therapy, or terminate it later, than patients with poorer score values. A later study on the same scale (Cromer and Hilsenroth 2010) confirmed the significant positive relationship between the CDPS total score and sessions attended, and further

demonstrated that of the nine individual items, the three items assessing reflective ability were statistically significantly related to the number of sessions attended.

The prediction by individual suitability items of psychotherapy outcome was reported for 7 scales (Husby et al. 1985, Husby 1985a, Barth et al. 1988a, Høglend et al. 1992a, b, 1993, 1994, Høglend 1993; Brodaty et al. 1982; Buckley et al. 1984; Piper et al. 1985, de Carufel and Piper 1988; Sigal et al. 1999; Myhr et al. 2007; Cromer and Hilsenroth 2010). In the study by Husby (1985b), the majority of 36 patients, all evaluated as being suitable for the assigned STPP – either STAPP or BDP – experienced improvement in their social functioning and self-esteem (78%), symptoms (67%), new learning (61%), and interpersonal relations (53%) by the 2-year follow-up after the end of treatment; all of these improvements persisted or intensified further during the 5-year follow-up. Higher, and thus better, values in flexibility of interaction with the interviewer, motivation, problem-solving capacity, and interpersonal relationships, but not in focality, were statistically significantly correlated with higher, and thus better, values in outcome measures both at the 2-year and 5-year follow-up, thus supporting psychotherapy outcome prediction by suitability measures (Husby et al. 1985, Husby 1985a). In line with Husby et al., Høglend et al. (1992a, 1994) also found that flexible interaction, one aspect of motivation (desire for change), and better quality of relationships, and, in contrast to Husby et al., focality but not problem-solving capacity correlated with a better dynamic outcome, reflecting new learning and self-understanding as well as changes in self-esteem, overall personality, problem-solving capacity, and interpersonal relations, at a 4-year, but not at a 2-year, follow-up. The best predictor of dynamic outcome was an interaction term between flexible interaction, motivation, and circumscribed focus which predicted both the two years' (N = 34) and four years' (N = 39) outcome (Høglend et al. 1992a). Quality of relationships was the only suitability item also predicting symptomatic change, at the 2-year follow-up (Høglend et al. 1992a, 1993, Høglend 1993). Furthermore, patients with a good level of interpersonal relationships benefited more from shorter, more focused treatment (Høglend et al. 1992a), whereas patients with a low or moderate level of interpersonal relationships and problem-solving capacity benefited more from a longer treatment (Høglend et al. 1993, 1994). In a study by Barth et al. (1988a; N = 44), assessing patient suitability using the same scale as Husby et al. (1985), 86% of the patients assigned to either STAPP, BDP, or FIAT approach attained substantial symptom relief and maintained it up to the 2-year follow-up; the patients in the FIAT group, however, reached their peak of change later than patients in STAPP and BDP, thus also implying that higher suitability may predict faster recovery. In another study on a different scale, Brodaty et al. (1982) found that the abilities to emote and accept interpretations correlated statistically significantly with global improvement at the end of dynamic problem-oriented therapy, and that ability to emote and psychological mindedness (overall and especially in women) correlated with a reduction in psychiatric symptoms at the 1-year follow-up in a

small sample of 18 patients; the other four suitability items on the scale did not correlate with the psychotherapy outcome. Buckley et al. (1984) found that 5 out of 50 observer-rated suitability items (undoing, rationalization, isolation, withdrawal, and independent personality), mostly related to defense mechanisms, correlated highly with 21 patients' short-term psychodynamic psychotherapy outcome, evaluated from multiple aspects: patients with less independency and a greater tendency to use such defenses appeared to improve the most. Similarly, another study of the same size (Piper et al. 1985) also found that the patient's defense style, and quality of relationships with important objects, predicted change in both patient- and therapist-rated target objectives and their severity as well as the overall usefulness of individual short-term psychodynamic psychotherapy, but the remaining 13 psychological suitability items (see Table 3) did not predict the psychotherapy outcome. A later study by de Carufel and Piper (1988) on the same patient sample and three other patient samples of a similar size, treated with group STPP, individual LTPP, and group LTPP, further revealed the suitability items capable of differentiating improvers from non-improvers within these therapy forms: quality of relationships predictively differentiated improvers from non-improvers in individual STPP, psychodynamic formulation, psychological mindedness, and type of relationships in group STPP, likability of the patients, motivation, and type of relationships in individual LTPP, and object choice and level of psychosexual conflict in group LTPP. A study by Myhr et al. (2007), applying the 9-item SSCT scale to a larger patient sample (N = 113), demonstrated that security operations, reflecting defensive information processing strategies, and awareness of emotions were the two scale items most strongly correlated with the symptom outcome of short-term cognitive therapy. Of the 9 individual items of the CDPS, on the other hand, only one of the three suitability items related to reflective ability (manifests insight) and the insight factor based on all three of them were statistically significantly associated with the STPP outcome (Cromer and Hilsenroth 2010; N = 71). Sigal et al. (1999) was the only study which failed to find correlations between any of the suitability items on their brief psychodynamic psychotherapy suitability checklist and the psychotherapy outcome.

The psychotherapy outcome prediction by overall suitability was reported for 4 scales (Brodaty et al. 1982; Alpher et al. 1990, Cromer and Hilsenroth 2010; Safran et al. 1993, Myhr et al. 2007; Jørgensen et al. 2000, Valbak et al. 2004). Brodaty et al. (1982) found that the sum of the 6 scale items predicted change in symptoms by the 1-year follow-up. Alpher et al. (1990) demonstrated that better CDPS score values were statistically significantly correlated with lesser symptoms, better functioning, and greater global change at the end of STPP in a sample of 25 patients. However, in a later study, Cromer and Hilsenroth (2010) failed to find a significant association between the CDPS score and any of the several patient- and therapist-rated outcome measures of symptoms and functioning in a larger sample of 71 patients. Safran et al. (1993; N = 42) studied the correlations between the mean

Patient suitability for short-term and long-term psychotherapy

SSCT score and multiple, both therapist- and patient-rated, outcome measures, including symptoms, coping styles, target complaints, and global success ratings, and found most of these (9 out of 12 outcome measures) to be statistically significant, even after adjustment for the baseline level of the outcome measure. A later study with a larger sample size (Myhr et al. 2007; N = 113) confirmed the findings on symptom outcome. Jørgensen et al. (2000) measured the correlations between the DAI ratings and general psychiatric symptoms, adjusted for the baseline symptom level, in a sample of 20 patients undergoing long-term psychodynamic psychotherapy (LTPP), but did not find that patients with high DAI ratings improved more than patients with low DAI ratings. A bigger study (N = 74) by Valbak et al. (2004) did not find the DAI score to correlate with a change in psychiatric symptoms either, but did find that it correlated with a change in social adjustment among patients mainly undergoing LTPP. Fisher et al. (1999) did not present quantitative results on the psychotherapy outcome prediction by the STS, but referred to another (still) unpublished study according to which the STS did predict the psychotherapy outcome, and could thus potentially be used to indicate whether certain types of psychotherapy procedures (e.g., abreactive, supportive) were likely to be more effective than others.

2.4.6 Summary of the findings on the validation of the psychotherapy suitability assessment scales

Relatively few studies have focused on the reliability and validity of the interview-based psychotherapy suitability assessment scales as an object of research: only 8 scales have been studied for reliability and 5 of them also for validity, although for an assessment scale to be considered to provide accurate inferences, it must be shown to be both reliable and valid. In addition, 3 scales were studied for their prediction of the psychotherapy outcome without first assessing their reliability or validity. Of the 5 scales studied for both reliability and validity, 4 have been shown to be both reliable and valid: STAPP, CDPS, SSCT, and STS. All of them have also been reported to be predictive of the psychotherapy outcome; unlike for other scales, for STS no results on its prediction were presented and the reference for the prediction results remains unpublished (Fisher et al. 1999).

Of the 4 validated scales, STAPP and CDPS were designed to evaluate suitability for short-term psychodynamic psychotherapy and SSCT for short-term cognitive therapy, whereas STS was considered applicable for the evaluation of suitability for different kinds of psychotherapies. STS, however, consists of over 200 items, rendering it highly time-consuming in comparison to STAPP, CDPS, and SSCT, which consist of less than 15 items. Moreover, the STS items assess a variety of patient characteristics, whereas STAPP, CDPS, and SSCT items focus on patients' psychological suitability. A 3-factor structure of the STS has been presented (Fisher et al. 1999) but no overall suitability score has been provided, unlike for STAPP,

CDPS, and SSCT, which would enable evaluation of the balance between patients' different psychological capacities. STAPP, CDPS, and SSCT cover several aspects of the nature of patients' problems, ego strength, and self-observing capacity, but none of them evaluates patients' self-concept or response to trial interpretation, despite these aspects having been considered very influential on psychotherapy suitability (Davanloo 1978). STAPP evaluates patients' intelligence, or in later modifications, problem-solving capacity, which is closely related to reflective ability. The validated scales cover different aspects of intrapsychic and interpersonal behavior, but focus on the evaluation of interpersonal relationships. The lack of items evaluating intelligence, coping styles and cognitive skills, defense styles, and personality traits is probably partly due to the existence of several well-validated scales specifically designed for the evaluation of these capacities. All validated suitability assessment scales were evaluated within one interview, although a suitability evaluation based on more than one interview is considered more reliable (see Section 2.2).

If only the 4 existing reliable and valid suitability assessment scales were used to evaluate the prediction of psychotherapy outcome by different aspects of psychological suitability, the relevance of all aspects could not be evaluated (see Table 1). As different studies also evaluate different dimensions of psychotherapy outcome (e.g., change in psychiatric symptoms, intrapsychic and interpersonal functioning, and self-esteem), replication of the findings could not necessarily be evaluated. More importantly, as the prediction by the validated scales and the suitability items included in them was only evaluated in relation to the outcome of a certain type of short-term therapy, their relevance in terms of referral to other types of short-term therapies or to long-term therapy could not be evaluated. For these reasons, a more thorough review of the psychological suitability measures – whether assessing one or more aspects of the patient's suitability and evaluated by interview or self-report – as predictors of psychotherapy outcome is presented in the following section.

2.5 Review of studies on suitability measures as predictors of psychotherapy outcome

To be able to refer a psychiatric patient to an optimal treatment, the referring physician should be able to anticipate the patient's likely outcome in therapies of different length and orientation. This requires knowledge of the patient's suitability for psychotherapy, i.e., of whether the patient has the necessary resources to undergo therapy and of which of the therapies, for which the patient has the resources, the patient is likely to benefit the most. Several aspects of the patient have been considered relevant in the evaluation of those resources, with the type, severity, and duration of the patient's disorder and psychological capacities being among the most important (see Section 2.2). To ascertain their relevance in the psychotherapy

process, an assessment must be performed of the ability of these aspects to predict the psychotherapy outcome. To evaluate their importance in psychotherapies of different length and duration, their ability to predict the outcome of different psychotherapies must be compared. The psychotherapy outcome prediction by the type, severity, and duration of the disorder has been widely studied in several recent reviews and meta-analyses (Leichsenring and Rabung 2008, 2011, Driessen et al. 2010, Leichsenring et al. 2013). The psychotherapy outcome prediction by psychological suitability has, however, only been summarized in one review ten years ago (Valbak 2004) although much literature on the topic has appeared since. Here, an up-to-date review of psychotherapy outcome prediction by psychological suitability measures is presented.

The psychological suitability selection criteria (Table 1) were originally proposed for separating patients treatable with short-term psychodynamic psychotherapies from those in need of long-term psychodynamic psychotherapies (see Section 2.2). Gradually, the development and application of these suitability selection criteria expanded to different types of psychotherapies. Accordingly, three major research questions can be formulated regarding both general and differential psychotherapy outcome prediction by psychological suitability measures:

1. a) Do psychological suitability measures predict the outcome of different short-term psychotherapy modalities, and b) does their prediction differ between different short-term psychotherapy modalities?
2. a) Do psychological suitability measures predict the outcome of different long-term psychotherapy modalities, and b) does their prediction differ between different long-term psychotherapy modalities?
3. Do psychological suitability measures predict the differential outcome between short-term and long-term psychotherapies?

Studies on one particular psychotherapy, which constitute most of the literature, may contribute to research questions on general psychotherapy outcome prediction by psychological suitability measures (1a and 2a). However, studies including several psychotherapies of different orientation and length are needed in order to answer research questions on differential psychotherapy outcome prediction by psychological suitability measures (1b, 2b, and 3). This review will cover the literature available for answering these three questions on psychological suitability measures, reflecting the suggested suitability selection criteria (Table 1), well-defined individual psychotherapies, and one major psychotherapy outcome area, i.e., psychiatric symptoms.

2.5.1 Selection of studies

The review covers published original quantitative studies on prediction by psychological patient suitability measures (Table 1) assessed prior to the start of the

therapy, either based on an interview or self-report, on changes in the psychiatric symptoms and functional ability of adult outpatients suffering mainly from Axis I disorders and treated with well-established short-term or long-term individual psychotherapy (Table 5). Studies were identified by checking the PubMed databases and examining the reference lists of identified articles. The search yielded a total of 54 original articles which fulfilled all of the selection criteria (Table 6). The articles were based on 41 separate datasets (Table 6). Of these, 38 included only short-term psychotherapies and 3 only long-term psychotherapies. Different short-term psychotherapy modalities were compared in 9 studies based on 6 datasets, whereas different long-term psychotherapy modalities were compared in only one study. No studies were published comparing the psychotherapy outcome prediction by psychological suitability measures in short-term versus long-term psychotherapy.

2.5.1.1 Data extraction

The following data were extracted from each study: the first author's name, year of publication, research site and duration (if available), study design, number of patients, patients' sex, age, and diagnoses on Axis I and Axis II, treatment forms, number of therapy sessions, training of therapists, psychological suitability measures and symptom measures used, duration of follow-up after the end of therapy, statistical methods used, potential confounding factors adjusted for in the analysis, and results on the strength of association between the pre-therapy psychological suitability measure and the post-therapy or post-follow-up symptom outcome and its significance (Table 6, Supplementary tables 1A-F). Wherever references for earlier studies on more detailed data description were provided, they were reported (Table 6) and and checked. Any information not available, in original studies or additional references provided, was marked as not reported (NR).

Table 5. Criteria for inclusion of studies for literature review on the suitability measures as predictors of psychotherapy outcome.

---

1. Type of study: Published original quantitative study

2. Predictor variable: Pre-therapy psychological patient suitability measure[1]

3. Outcome variable: Psychiatric symptoms and functional ability

4. Study design: Randomized trial or cohort study

5. Psychotherapy: Well-established short-term or long-term individual psychotherapy

6. Main diagnosis: Axis I disorder

7. Patients: Outpatients

8. Age: Adults

9. Result: Strength of association or significance of association reported

---

[1] See Table 1.

Table 6. Description of the datasets and individual studies on suitability measures as predictors of individual short-term and long-term psychotherapy outcome.[1]

	Data & Location[2]	Author[3]	Diagnosis (%)[4]	N[5]	Males (%)	Age[6]	Treatment[7]		Sessions[8]	Therapist training[9]	Predictor[10]	Outcome[11]	Follow-up[12]
Short-term therapies													
1	Department of Clinical Psychology, Academic hospital Groningen, Netherlands	**Emmelkamp 1980**	AXIS I 100% ANX 100%[13]	17	35	37 (18-56)	BT		4 (4-4)	ST	D5:P12:1,2	V1:1 V2:1a-f, 2a-f,3a,b,4	0,1
2	Yale University, New Haven and Tufts Medical School Boston, USA	**Zuckerman et al. 1980**	AXIS I 100% DEP 100%[13]	81	15	34 (18-65)[13]	IPT (N=NR) MED (N=NR) IPT+MED (N=NR) NTC (N=NR)		13[14] (2-16)[14]	NR	D5:P9:4 P11:28, 29, 37	V1:2,3	0,12
3	10 urban general practices Sydney, Australia		AXIS I NR[15]	128[16]	NR	NR (18-65)[13]	PP(I) (N=51) OTHER (N=77)[17]		8[18,19] (NR)	NR			
		Brodaty et al. 1982	AXIS I NR[15]	18[20]	44	49 (23-64)	PP(I)		8[18,19] (5-NR)[20]	T	D1:P1:1[21] D2:P2:1[21] D3:P5:1[21] P6:1[21] P7:1[21] D5:P11:1[21] D6:P13:1[21]	V3:1	0,12
4	University clinical research center San Francisco, USA	**Horowitz et al. 1984**	AXIS I 100% DEP 13% ANX 33% COM I 4%	52[22]	4	39 (21-72)	PP(I)		12[19] (NR)	T	D2:P4:1 D3:P7:2	V3:1	0
5	The Oslo Group Oslo, Norway 1973-1976	**Husby et al. 1985, Husby 1985a**	AXIS I NR[23]	39[24]	49	28[25] (21-35)	PP(I)		20[25] (4-39)	T	D1:P1:2,3 D2:P3:1 D3:P7:3 D4:P8:1 D5:P12:3	V3:2	24,60

#	Setting	Author, year	Diagnosis	N	Age	Age (range)	Treatment (N)	Duration (weeks)	T/NR	Code	V	
6	Washington University Outpatient Psychiatry Clinic Washington, USA	Murphy et al. 1984	AXIS I 100% DEP 100%[13]	70[16]	NR	NR (18-60)[13]	CT (N=19), MED (N=16), CT+MED (N=17), CT+PLA (N=18)	12[14,19] (NR)	NR	D5:P9:1,10-13	V1:4	0
		Simons et al. 1985	AXIS I 100% DEP100%[13]	35[24]	NR	NR (18-60)[13]	CT (N=19), MED (N=16)	12[14,19] (NR)	NR	D6:P13:2	V3:1,3	0
7	University affiliated psychotherapy research clinic USA	**Alpher et al. 1990**	AXIS I 100% DEP 76% ANX 4% AXIS II 72%	25[24]	NR	NR (21-65)	PP(I)	23 (NR-25)	T			0
8	Center for Cognitive Therapy, University of Pennsylvania Medical School USA 1986-1988	**Haaga et al. 1991**	AXIS I 100% DEP 72%[13] ANX 28%[13]	106[20,24]	NR	NR	CT	21 (6-NR)[20]	NR	D4:P8:3,4	V1:4 V2:5	0
9	Mood Disorders Program, University of Texas Southwestern Medical Centre Dallas, USA 1983-1988	**Jarrett et al. 1991a**	AXIS I 100% DEP 100%[13]	37[20,24]	35	37 (NR)	CT	18 (6-20)[20]	T	D5:P9:1,16	V1:3,4	0
		Jarrett et al. 1991b	AXIS I 100% DEP 100%[13]	37[20,24]	35	37 (NR)	CT	18 (6-20)[20]	T	D5:P9:13	V1:3,4	0
10	National Institute of Mental Health Treatment of Depression Collaborative Research Program (NIMH-TDCRP) Washington, Pittsburgh and Oklahoma, USA	Elkin et al. 1985, 1989 Imber et al. 1990	AXIS I 100% DEP 100%[13]	239[22]	30	35 (21-60)[13]	CBT (N=59), IPT (N=61), MED+CM (N=57), PLA+CM (N=62)	13 (NR)	T			
		Sotsky et al. 1991	AXIS I 100% DEP100%[13]	162[20]	33	35 (21-60)[13]	CBT (N=37), IPT (N=47), MED+CM (N=37), PLA+CM (N=34)	16 (12-NR)[20]	T	D3:P7:10 D5:P9:9 P12:10	V1:3,4	0
		Blatt et al. 1995	AXIS I 100% DEP 100%[13]	155	30	NR (21-60)[13]	CBT (N=NR), IPT (N=NR), MED+CM (N=NR), PLA+CM (N=NR)	NR	T	D5:P9:2,3	V1:3,4 V3:1,4	0

Table 6. Description of the datasets and individual studies on suitability measures as predictors of individual short-term and long-term psychotherapy outcome.[1]

Data & Location[2]	Author[3]	Diagnosis (%)[4]	N[5]	Males (%)	Age[6]	Treatment[7]	Sessions[8]	Therapist training[9]	Predictor[10]	Outcome[11]	Follow-up[12]
Short-term therapies											
11 Oslo, Norway 1978-1987		AXIS I NR	47[16]	NR	NR	PP(I)	NR	T			
	Høglend et al. 1992a	AXIS I 65% DEP 13% ANX 25% AXIS II 33%	40[22]	35	32 (20-53)	PP(I)	28 (9-53)	T	D5:P12:4	V3:4	24,48
	Høglend 1993	AXIS I 65% DEP 12% ANX 23% AXIS II 33%	43[20]	33	32 (20-53)	PP(I)	28 (9-53)	T	D1:P1:4 D2:P3:2 D3:P7:6 D5:P12:4	V3:4	24,48
	Høglend et al. 1993	AXIS I 65% DEP 12% ANX 23% AXIS II 33%	43[20]	33	32 (20-53)	PP(I)	28 (9-53)	T	D5:P12:4	V3:4	24,48
	Høglend et al. 1994	AXIS I 65% DEP 12% ANX 23% AXIS II 33%	43[20]	33	32 (20-53)	PP(I)	26[26] (9-53)	T	D3: P8:2	V3:4	24,48
12 Clarke Institute of Psychiatry Toronto, Canada		AXIS I 100%[27]	64[16]	53	38 (23-62)	CBT (N=42) OTHER (N=22)[28]	NR	NR			
	Safran et al. 1993	AXIS I 100%[27]	42[22]	52	37 (23-62)	CBT	20[19] (NR)	NR	D6:P13:3	V1:4-6 V2:6 V3:1	0
13 University outpatient clinic Nijmegen, Netherlands	**Keijsers et al. 1994a**	AXIS I 100% ANX 100%[13]	60[29]	27	36 (18-59)	BT	12[19] (NR)	ST	D3:P7:7	V2:7a-d, 8	0,2
14 University outpatient clinic Nijmegen, Netherlands	**Keijsers et al. 1994b**	AXIS I 100% ANX 100%[13]	40[29]	45	35 (NR)	BT	34 (NR)	ST	D3:P7:7	V2:9,10	0

No.	Setting	Study	Diagnoses	N		Age	Treatment		Duration	Type			
15	Western Psychiatric Institute and Clinic Pittsburgh, USA	**Simons et al. 1995**	AXIS I 100% DEP100%[13]	53[20,29]	23	38 (NR)	CBT		19 (NR)	T	D5:P9:1	V1:3,4	0
16	Mood disorders module of the western psychiatric clinic, University of Pittsburgh Pittsburgh, USA	**Spangler et al. 1997**	AXIS I 100% DEP 100%[13]	59	25	38 (NR)	CBT		20[19] (NR)	T	D5:P9:4,5, 17,18	V1:3,4	0
17	York University Toronto, Canada 1991-1993	**Paivio & Bahr 1998**	AXIS I 100%[30]	33[22]	36	42 (23-63)	ET		12 (8-16)	T	D5:P12:16-25	V3:1	0
18	University of Alberta Hospital Site Edmonton, Canada 1993-1996	**Piper et al. 1998**	AXIS I 73% DEP 55% ANX 6% AXIS II 60% COM II 47%	144[29]	39	34 (18-62)	PP(I) (N=72) PP(S) (N=72)		18 (NR)	T	D3:P6:2 D5:P12:5	V1:4 V2:11 V3:1	0
19	University affiliated clinic Montreal, Canada	**Sigal et al. 1999**	AXIS I 100% DEP 20% ANX 10%	43[29]	25	35 (21-63)	PP(I)		12 (12-12)	T, ST	D1:P1:5[21] D2:P2:2,3[21] D3:P7:11-13[21] D5:P11:2[21] P12:6,7[21]	V3:5,6	0
20	Vanderbilt II Psychotherapy Research Project, Vanderbilt University Nashville, USA	Henry et al. 1993	AXIS I 87% AXIS II 67%	84[22]	23	41 (24-64)	PP(I)		21[31] (5-25)	T, ST, NT			
		Hilliard et al. 2000	AXIS I NR	64[32]	NR	NR	PP(I)		NR (NR-25)	T, NT[32]	D5:P12:8	V3:1	0
21	National Health Service and University Research Clinic UK		AXIS I 100% DEP 100%[13]	110[16]	27[33]	35[33] (NR)	CT		NR (12-20)[19]	T			
		Hardy et al. 2001	AXIS I 100% DEP 100%[13]	24[20]	29	35 (20-56)	CT		NR (12-20)[20]	T	D5:P9:6-8 P12:29,30	V1:4	0
		Saatsi et al. 2007	AXIS I 100% DEP 100%[13]	106[22]	27[33]	35[33] (NR)	CT		12[34] (NR-20)	T	D5:P12:31-33	V1:4	0

Table 6. Description of the datasets and individual studies on suitability measures as predictors of individual short-term and long-term psychotherapy outcome.[1]

	Data & Location[2]	Author[3]	Diagnosis (%)[4]	N[5]	Males (%)	Age[6]	Treatment[7]	Sessions[8]	Therapist training[9]	Predictor[10]	Outcome[11]	Follow-up[12]
Short-term therapies												
22	Norwegian Multisite Study of Process and Outcome in Psychotherapy (NMSPOP) Norway 1995	**Hersoug et al. 2002**	AXIS I NR, ANX 67%, AXIS II 65%	43	14	36 (NR)	PP(I)	32 (1-40)	T	D5:P10:1,2	V3:1,4	0
23	The University of Texas Southwestern Medical Center, Psychosocial Research and Depression Clinic Dallas, USA	Jarrett et al. 2001	AXIS I 100%[13], DEP 100%[35], AXIS II 70%[35]	156[22]	NR	NR	CT	20[19] (NR)	T			0
		Clark et al. 2003a	AXIS I 100%, DEP 100%[13], AXIS II 70%[35]	108[24]	NR	NR	CT	20[19] (NR)	T	D5:P11:3-17[21]	V1:7	0
		Clark et al. 2003b	AXIS I 100%, DEP 100%[13], AXIS II 70%[35]	100[24]	NR	NR	CT	20[19] (NR)	T	D5:P9:1,16, 17,20[21] P11:3-8[21] P12:9,15[21]	V1:7	0
24	Department of Psychiatry, University of Toronto Toronto, Canada 2001-2004		AXIS I 100%, DEP 100%[13]	137[16]	NR	40 (18-60)[13]	CBT (N=43), IPT (N=39), MED+CM (N=55)	NR (16-20)[19]	T, ST		V1:3,4	0
		McBride et al. 2006	AXIS I 100%, DEP 100%[13], ANX 13%, COM I 20%, AXIS II 5%	56[29]	27	41 (18-60)[13]	CBT (N=29), IPT (N=27)	17 (16-20)[19]	T, ST	D5:P12:34,35	V1:3,4	0
		Zuroff et al. 2007	AXIS I 100%, DEP 100%[13], ANX 13%, COM I 18%, AXIS II 8%	95	29	42 (NR)	CBT (N=37), IPT (N=35), MED+CM (N=30)	NR (16-20)[19]	T, ST	D3:P7:8,9[21]	V1:3,4	0

#	Setting	Study	Diagnosis	N	Age	Age mean (range)	Treatment	N per group	Sessions mean (range)	Therapy	D codes	V codes	
		Marshall et al. 2008	AXIS I 100%[13] DEP 100%[13] ANX 13% COM I 19% AXIS II 12%	102[24]	31	NR (18-60)[13]	CBT IPT MED+CM	(N=37) (N=35) (N=30)	NR (16-20)[19]	T, ST	D5:P11:25,26	V1:3	0
25	Netherlands	Blom et al. 2007a **Blom et al. 2007b**	AXIS I 100%[13] DEP 100%[13]	193	36	40 (18-NR)[13]	IPT MED IPT+MED IPT+PLA	(N=50) (N=47) (N=49) (N=47)	12 (12-12)	T	D5 P11:31-35	V1:3	0
26	The Christchurch Psychotherapy for Depression Study (CPDS), University Clinical Research Unit New Zealand 1998-2003	Luty et al. 2007	AXIS I 100%[13] DEP 100%[13]	177[16]	28	35 (18-NR)[13]	CBT IPT	(N=86) (N=91)	16[19] (NR)	T		V1:3	0
		Joyce et al. 2007	AXIS I 100%[13] DEP 100%[13] AXIS II 54%[35]	167[24]	27	36 (18-NR)[13]	CBT IPT	(N=80) (N=87)	16[19] (NR)	T	D5:P11:18-24	V1:8	0
		Carter et al. 2011	AXIS I 100%[13] DEP 100%[13]	177	28	35 (18-NR)[13]	CBT IPT	(N=86) (N=91)	16[19] (NR)	T	D5:P9:1 P12:11-14	V1:8	0
27	The McGill University Health Centre, Cognitive Therapy Unit Montreal, Canada 2001-2006	**Myhr et al. 2007**	AXIS I 80% DEP 24% ANX 56% COM I 45% COM II 16%	113[24]	27	36 (NR)	CBT		19 (3-65)	T, ST	D1:P1:6 D2:P2:4 P3:3 D3:P6:8 P7:14-16 D5:P10:16 D6:P13:3	V3:7	0
28	Outpatient clinic for anxiety disorders and university outpatient clinic Netherlands	**Kampman et al. 2008**	AXIS I 100%[13] ANX 100%[13] AXIS II 37%	129	34	36 (18-65)[13]	CBT		15[19] (NR)	T, ST	D3:P7:7	V2:7a-d,10, 12,13a,b	0,1
29	2 outpatient facilities of the Jellinek Mentrum Mental Health Care psychiatric teaching hospital Amsterdam, Netherlands	Dekker et al. 2008	AXIS I 100%[13] DEP 100%[13]	166[22,36]	26	NR (18-65)[13]	PP(S) MED	(N=99)[36] (N=67)[36]	16[19] (NR)	T	D5:P12:37-46	V1:3	0
		Van et al. 2008	AXIS I 100%[13] DEP 100%[13]	81[22,24]	21	NR (20-65)[13,26]	PP(S)		16[19] (NR)	T		V1:3	0
		Van et al. 2009	AXIS I 100%[13] DEP 100%[13]	81[22,24]	21	NR (20-65)[13,26]	PP(S)		16[19] (NR)	T	D5:P10:2-15	V1:3	0

Table 6. Description of the datasets and individual studies on suitability measures as predictors of individual short-term and long-term psychotherapy outcome.[1]

Data & Location[2]	Author[3]	Diagnosis (%)[4]	N[5]	Males (%)	Age[6]	Treatment[7]	Sessions[8]	Therapist training[9]	Predictor[10]	Outcome[11]	Follow-up[12]
Short-term therapies											
30 University of Pennsylvania, Philadelphia and Vanderbilt University, Nashville USA 1998-2003	DeRubeis et al. 2005	AXIS I 100% DEP 100%[13] ANX 53% COM I 72% COM II 48%	240	41	40 (18-70)	CT (N=60) MED (N=120) PLA (N=60)	NR (20-28)[19]	T			
	Fournier et al. 2009	AXIS I 100% DEP 100%[13]	180	NR	NR (18-70)	CT (N=60) MED (N=120)	16[14] (NR)	T	D4:P8:5 D5:P9:2,3, 12,15,21	V1:3	0
31 USA	Resick et al. 2002	AXIS I 100% ANX 100%[13]	171	0	NR	BT (N=62) CBT (N=62) WL (N=47)		T			
	Rizvi et al. 2009	AXIS I 100% ANX 100%[13]	145	0	32 (18-70)	BT (N=73) CBT (N=72)	13[19] (1-NR)	T	D4:P8:6 D5:P11:36	V2:14	0
32 The Research Centre of the Social Insurance Institution Turku, Finland 2000-2004		AXIS I 100% DEP 100%[13] AXIS II 0%	50[16]	32	43 (20-60)[13]	PP(I) (N=25) MED (N=25)	16[19] (NR)	T			0,8
	Kronström et al. 2009	AXIS I 100% DEP 100%[13] AXIS II 0%	35[24]	NR	NR (20-60)[13]	PP(I) (N=19) MED (N=17)	16[19] (NR)	T	D3:P6:3 D5:P10:3-5	V1:3	0,8
	Kronström et al. 2011	AXIS I 100% DEP 100%[13] AXIS II 0%	35[24]	NR	NR (20-60)[13]	PP(I) (N=19) MED (N=17)	16[19] (NR)	T	D5:P11:18-24	V1:3	0
33 Outpatient mood disorders clinic of a large university-affiliated psychiatric hospital Canada	**Cromer & Hilsenroth 2010**	AXIS I 100% DEP 54% ANX 14% AXIS II 54%	71[22]	27	30 (NR)	PP(I)	27 (NR)	ST	D2:P2:5-8 P3:4 D3:P6:4-7 D5:P11:27 P12:47,48 D6:P13:2	V3:3,8	0

No.	Setting / Country	Author, Year	Diagnosis	N	%	Age	Treatment	N	Sessions	Format	D	V	T
34	Canada	**McBride et al. 2010**	AXIS I 100% DEP 100%[13] ANX 16% COM I 39%	74[37]	26	40 (18-68)[13]	IPT		16[19] (NR)	T, ST	D3:P7.8,9[21]	V1:4	0
35	Massachusetts General Hospital OCD Clinic, University of Virginia Boston, USA	**Steketee et al. 2011**	AXIS I 100% ANX 100%[13] DEP 36% COM I 49%	39	46	33 (18-NR)[13]	CT		22[19] (NR)	ST	D3:P7:4[21]	V2:15	0,12
36	University training clinic USA	**Lewis et al. 2012**	AXIS I 92% DEP 56% ANX 32%	173	34	28 (18-64)	CBT		NR	ST	D3:P7:5A-D D5:P10:1,11	V1:4	0
37	Continuation Phase Cognitive Therapy Relapse Prevention (C-CT-RP) Trial, The University of Pittsburgh Medical Center and The University of Texas Southwestern Medical Center USA	**Renner et al. 2012**	AXIS I 100% DEP 100%[13]	523[22]	33	42 (18-70)[13]	CT		(NR) (16-20)[19]	NR	D5:P12:26-28	V1:3	0
38	USA		AXIS I 100% ANX 100%[13]	147[16]	43	NR (18-60)[13]	ACT (N=66) CBT (N=81)		12[19] (NR)	ST			0,6,12
		Wolitzky-Taylor et al. 2012	AXIS I 100% ANX 100%[13]	87[20]	53	38 (18-60)[13]	ACT (N=38) CBT (N=49)		12 (12-12)	ST	D5:P11:30	V2:16	0,6,12

Table 6. Description of the datasets and individual studies on suitability measures as predictors of individual short-term and long-term psychotherapy outcome.[1]

Long-term therapies

	Data & Location[2]	Author[3]	Diagnosis (%)[4]	N[5]	Males (%)	Age[6]	Treatment[7]	Sessions[8]	Therapist training[9]	Predictor[10]	Outcome[11]	Follow-up[12]
39	Therapeutic Alliance Research Project, Psychiatric University Hospital Aarhus, Denmark		AXIS I NR, AXIS II 67%	67	28	32 (20-50)	PP	NR	T, ST			
		Jørgensen et al. 2000	AXIS I NR, AXIS II 80%	20	35	30 (22-46)	PP	69 (20-135)	T, ST	D4:P12:49,50 D5:P13:4	V3:1	0
40	Transparency and Outcome Orientation in Outpatient Psychotherapy (TRANS-OP), Centre for Psychotherapy Research Stuttgart, Germany 1998-2002		NR	780	NR	NR	CBT (N=248) PP (N=397) PA (N=135)	25-50[19] 50-80[19] ≤ 300[19]	T			
		Puschner et al. 2004	AXIS I NR DEP 47% AXIS II 5%	622	45	44 (NR)	CBT (N=NR) PP (N=NR) PA (N=NR)	25-50[19] 50-80[19] ≤ 300[19]	T	D4:P12:24,25	V3:9	0
41	Norwegian Multisite Study of Process and Outcome in Psychotherapy (NMSPOP) Norway 1995	**Solbakken et al. 2012**	AXIS I NR DEP 50% ANX 35% AXIS II 51%	153	29	35 (18-65)	PP	72 (2-364)	T	D2:P2:9	V3:1	6,12,24

NR = not reported

[1] See Appendix 2 for detailed description of terminology and definitions.
[2] Number of separate datasets.
[3] Bolded studies fulfilled the selection criteria (see Table 5), other studies provided additional information on the data used in the selected studies. If the study sample in the suitability study was a subset of a larger database, characteristics of that database are also given.
[4] AXIS I % = Proportion of patients with Axis I disorders, DEP % = Proportion of patients with depressive disorder, ANX % = Proportion of patients with anxiety disorder, COM I % = Proportion of patients with comorbid Axis I disorders, AXIS II % = Proportion of patients with Axis II disorders (i.e., personality disorders), COM II % = Proportion of patients with comorbid Axis I and Axis II disorders.
[5] Number of patients and related population characteristics in the original articles were reported at different phases of the study (i.e., patients randomized/referred to treatments, patients who began treatment, patients who completed treatment, patients who were analyzed). Patient characteristics for the patients used in the analysis is reported where possible.
[6] Mean age and age range unless otherwise mentioned.
[7] ACT = Acceptance and Commitment Therapy, BT = Behavioral Therapy, CBT = Cognitive-Behavioral Therapy, CT = Cognitive Therapy, CM = Clinical Management, ET = Experiential Therapy, IPT = Interpersonal Therapy, MED = Medication/Pharmacotherapy, NTC = Nonscheduled Treatment Control, PA = Psychoanalysis, PP = Psychodynamic Psychotherapy, PP (I) = PP (Interpretative),

Patient suitability for short-term and long-term psychotherapy

PP (S) = PP (Supportive), PLA = Placebo, WL = Waiting List.

[8] Mean number and range of therapy sessions unless otherwise mentioned.

[9] Therapists were categorized into three groups according to their training and experience: NT = Untrained, ST = Semi-trained, T = Trained.

[10] *D1. Nature of problems:* P1. Focus

D2. Ego strength: P2. Modulation of affects, P3. Flexibility of interaction, P4. Self-concept in relation to ego ideal

D3. Self-observing capacity: P5.Response to trial interpretation, P6. Reflective ability, P7. Motivation

D4. Intelligence: P8. Intelligence

D5. Intrapsychic and interpersonal behavior: P9. Coping styles and cognitive skills, P10. Defense styles, P11. Personality traits, P12. Interpersonal relationships

D6. Overall suitability: P13. Overall suitability

More detailed description of the individual suitability measures is given in Appendix 2.

[11] *Psychiatric symptoms and functional ability:* V1. Depressive symptoms, V2. Anxiety symptoms, V3. Global symptoms and functioning

More detailed description of the individual outcome measures is given in Appendix 2.

[12] Length of post-treatment follow-up (months from the end of the treatment).

[13] Inclusion criteria.

[14] Weeks.

[15] Persistently symptomatic emotionally disordered patients.

[16] Patients who were randomized / referred to treatment.

[17] Other treatment than psychotherapy.

[18] 30-minute therapy sessions.

[19] Intended number of therapy sessions

[20] Completion of a certain number of therapy sessions required:

Broday et al. 1982: 5/8; Haaga et al. 1991: 6; Jarrett et al. 1991a, b: 6/20; Sotsky et al. 1991: 12/16-20; Høglend 1993, Høglend et al. 1994: 7; Simons et al. 1995: full 16 weeks (20 sessions);

Hardy et al. 2001: 12/12-20; Wolitzky-Taylor et al. 2012: 12 / 12

[21] Suitability measure evaluated soon after the start of treatment:

Broday et al. 1982: At the end of first or second therapy session; Sigal et al. 1999: After the first session; Clark et al. 2003a, b: At the second session; Zuroff et al. 2007: At the third session;

McBride et al. 2010: At the third session; Steketee et al. 2011: After session 4 or 6.

[22] Patients who began treatment.

[23] All patients were evaluated to have a neurotic personality structure.

[24] Patients for whom complete pre and post treatment data on the suitability and outcome measures used was available.

[25] Median.

[26] The reported information differs from information reported in prior studies for the same study population.

[27] Depressive and anxiety based disorders.

[28] E.g., long-term therapy and/or pharmacotherapy.

[20] Patients who completed therapy.

[30] Reported problems with depression and anxiety.

[31] Disregarding 4 patients who dropped out prior 5th therapy session (i.e., N=80).

[32] In the original study (Henry et al. 1993) each therapist saw 2 patients prior their training in STPP (Cohort 1), 1 patient during training (Cohort 2), and 2 patients after training (Cohort 3). This study was based on Cohorts 1 and 3 and 64 patients.

[33] Based on 106 patients who began treatment.

[34] Based on 97 patients who completed both pre and post treatment BDI.

[35] Based on patient self-report measure.

[36] Of the 204 eligible patients, 141 were randomized whereas 63 patients chose their treatment (40 PP(S), 23 MED). Of the patients randomized, 59 began PP(S) and 44 began MED.

[37] Subset of a larger naturalistic database of patients seen in the clinic (the details of which are not reported).

2.5.1.2 Standardization

Because the studies included in the literature review reported variables of interest using different terms and applied different assessment methods, the following standardization and grouping were carried out (Appendix 2).

The patient population for which the baseline characteristics were described varied from one study to another: patients randomized or referred for treatment, patients who had begun their treatment, patients who had completed the treatment or a certain minimum amount of therapy sessions, or patients for whom all of the relevant suitability and outcome assessments were available (i.e., the patient population used in the analysis). The number of patients for whom the baseline characteristics were described is given in the description of the study populations (Table 6) and the number of patients used in the analysis is given in the description of the results (Supplementary tables 1A-F). The baseline characteristics of the patients were reported in the following way: percentage of males, mean value and range of age, and percentage of patients diagnosed with Axis I disorders, with the percentage of patients diagnosed with depressive and anxiety disorders and comorbid Axis I disorders being reported separately (if available), as well as Axis II disorders, either as a main or comorbid disorder (Table 6).

Regarding the treatment provided, different treatment groups and the number of patients in each of them, the mean number and range of therapy sessions attended, the duration of patients' follow-up after the end of treatment (in months), and the training of the treating therapists were reported. A variety of different names and abbreviations for the same therapy form, psychodynamic psychotherapy in particular, were used in the original articles, and the therapy names and their abbreviations were therefore standardized (Appendix 2). All of the treatments provided in each study were reported in the description of the study populations (Table 6), including pharmacotherapy, treatment as usual, and placebo, but the results of symptom prediction by psychological suitability measures were only reported for well-defined individual short-term and long-term psychotherapies (Supplementary tables 1A-F), in line with the aim of this review. The training of the therapists was categorized into three groups: therapists adequately trained and experienced in the therapy they provided in the study, therapists under training or supervision for the therapy they provided in the study, and untrained therapists with potential training and/or experience in other therapy forms but not the one they provided in the study.

A large variety of variables describing different aspects of patients' psychological suitability were used as predictors in the original studies. They were grouped under the 13 previously introduced suitability predictor categories which were further placed under the 6 suitability domains, according to their clinically relevant conceptual scope (Table 1, Appendix 2). Although most of the variables could be clearly placed under one particular suitability category, some of the variables can be considered to represent or be related to more than one suitability

category (e.g., variables related to flexibility of interaction between the patient and the therapist, interpersonal relationships, and certain personality aspects such as collaborativeness; see Appendix 2 for the chosen classification of the original suitability measures). This is natural, as the description of psychotherapy suitability consists of several closely intertwined concepts and measures. Furthermore, even if different studies were intended to measure the same aspect of suitability and its role in the psychotherapy outcome prediction, most of them introduced and reported their own suitability measure. With the exception of the few validated suitability assessment scales and the suitability items included in them (see Section 2.4), these measures were often unvalidated. Only a few well-known and validated suitability measures (e.g., Dysfunctional Attitudes Scale (DAS; Weissman 1979)) were applied in more than one study. In addition to the suitability domain and category under which the individual suitability measure was classified, the type of the measure, i.e., whether a single variable (e.g., a single question or an item of a suitability scale) or a summary variable (e.g., a factor or a summary score based on several questions or items) was also reported (Supplementary tables 1A-F, Appendix 2).

Psychotherapy outcome was evaluated by assessing the change in psychiatric symptoms and the symptom measures were grouped into measures of depressive symptoms, anxiety symptoms, and global symptoms and functioning. The outcome type used in the analysis was categorized in the following way: 1) continuous post-therapy or post-follow-up value of the outcome measure, 2) binary post-therapy or post-follow-up value of the outcome measure (e.g., remission defined around a certain symptom value), 3) difference in the outcome measure between pre-therapy and post-therapy or post-follow-up assessments, 4) residual change score (i.e., the difference between the actual post-therapy or post-follow-up value of the outcome measure and the expected post-therapy or post-follow-up value of the outcome measure, predicted based on the regression of the pre-therapy outcome measure on post-therapy or post-follow-up outcome measure), 5) percentual improvement (i.e., the difference in the outcome measure between the pre-therapy and post-therapy or post-follow-up assessments divided by the pre-therapy value of the outcome measure), and 6) Reliable Change Index (i.e., the difference in the outcome measure between the pre-therapy and post-therapy or post-follow-up assessments divided by the standard error of the difference). The statistical method applied, whether correlation or modeling, and the potential confounding factors adjusted for in the analysis were reported. Adjustment of the baseline level of the symptom measure studied was of particular interest as pre- and post-therapy symptom measurements are likely to be related and an independent association between a suitability predictor and symptom outcome can be concluded to exist only if the baseline symptom level (and other factors potentially confounding the association) has been adjusted for. The level of adjustment for the symptom measure at baseline was categorized into three groups: no adjustment (i.e., use of one of the first three outcome types and no adjustment for the baseline symptom level), standardization

(i.e., use of one of the last three outcome types which have been standardized with respect to the baseline symptom level), and adjustment (i.e., use of one of the first three outcome types and adjustment for the baseline symptom level by including the symptom measure at baseline in the model). Adjustment by modeling is considered superior to standardization, and standardization to no adjustment.

2.5.2 Overview of the selected studies

2.5.2.1 Study characteristics

The first suitability studies fulfilling the selection criteria (Table 5) were conducted in the early 1980s, and from the beginning of the 1990s the number of suitability studies grew steadily (Table 6). The majority of the 38 short-term psychotherapy studies were carried out in the USA (45%) and Canada (21%) (Table 7). The European studies were carried out in the Netherlands (16%), Norway (8%), UK (3%), and Finland (3%). In addition, one study was conducted in Australia and one in New Zealand. The three long-term psychotherapy studies were all carried out in Europe, in Norway, Denmark, and Germany, in the 21st century.

The great majority (63%) of the short-term psychotherapy suitability studies were cohort studies exploring suitability for a particular type of short-term psychotherapy. One third of these studies were on interpretative psychodynamic psychotherapy (PP), for which the suitability criteria was first developed, but studies on cognitive therapy (CT; 16%) and cognitive-behavioral therapy (CBT; 13%) were also common. Of the randomized clinical trials (RCT) on short-term psychotherapies (37%), 21% compared short-term psychotherapy and another treatment form, most often some type of pharmacotherapy or combination of psychotherapy and pharmacotherapy, and 16% two different types of short-term therapies (possibly also with pharmacotherapy), with comparisons between CBT and interpersonal therapy (IPT) being most common (8%). One study compared prediction by psychological suitability between interpretative and supportive forms of psychodynamic psychotherapy, but no comparisons between short-term psychodynamic psychotherapy and other short-term psychotherapy modalities have been carried out, although most psychological suitability measures have been developed within a psychodynamic frame. Of the three studies investigating long-term psychotherapies, two were cohort studies on PP and one a cohort study including PP and CBT. No studies comparing the prediction of the outcome of short-term and long-term psychotherapies by psychological suitability measures have previously been carried out, despite the fact that the suitability criteria were developed for separating out suitability for short-term and long-term psychotherapy.

2.5.2.2 Patient characteristics

The number of patients undergoing a particular short-term psychotherapy, for whom the patient characteristics were reported, varied from 17 to 523, the median number

Patient suitability for short-term and long-term psychotherapy

being 52 (Table 7). The median number of patients used in the analysis, for whom the necessary suitability and symptom measurements before and after the psychotherapy were available, was 50. In only 5 short-term psychotherapy datasets (Clark et al. 2003a, b; Myhr et al. 2007; Kampman et al. 2008; Lewis et al. 2012; Renner et al. 2012) were the analyses based on more than 100 patients; these were all studies on either CT or CBT and within the last decade (Supplementary tables 1A-F). The largest study on PP involved 72 patients, of whom 64 were used in the analyses (Piper et al. 1998). The two long-term psychotherapy studies reporting the number of patients analyzed were based on 16 (Jørgensen et al. 2000) and 153 (Solbakken et al. 2012) patients who underwent PP. In the third long-term psychotherapy study (Puschner et al. 2004), 397 patients initially started PP and 248 patients CBT, of whom a total of 80% returned the baseline questionnaire and gave written consent to their participation; the number of patients on which the analyses were based was not, however, reported.

The age of adult outpatients recruited for short-term psychotherapies varied between 18-72 and to long-term psychotherapies between 18-65, the median mean age in short-term psychotherapies being 36 and in long-term psychotherapies 35 years (Tables 6 and 7). Three short-term psychotherapy studies (Haaga et al. 1991, Clark et al. 2003a, b, Fournier et al. 2009) did not report the mean age or age range of the patients (Table 6). In half of the 31 short-term psychotherapy datasets reporting the sex of the patients, less than one third (median 29%) of the patients were male (range 0%-53%); in long-term psychotherapy datasets the median percentage of males was 35 (range 29%-45%).

Of the 35 short-term psychotherapy datasets explicitly reporting the percentage of patients with Axis I diagnosis (range 65%-100%), in 30 all patients had Axis I diagnosis, with depressive and anxiety disorders being the most common diagnoses (Tables 6 and 7). Primary depressive disorder was a selection criterion in 16 datasets and primary anxiety disorder in 7 datasets; in the remaining datasets, depressive disorder was the prevalent disorder in 6 and anxiety disorder in 4. Five datasets (Brodaty et al. 1982; Husby et al. 1985, Husby 1985a; Safran et al. 1993; Paivio and Bahr 1998; Hilliard et al. 2000) did not provide details on diagnoses but indicated that the patients were persistently symptomatic and suffering from neurotic, mainly depressive and anxiety, disorders. In 13 short-term psychotherapy datasets reporting Axis II diagnoses, such disorders were present in a median of 54% of the patients (range 0%-72%). All three long-term psychotherapy studies comprised patients suffering from neurotic, mainly depressive disorders (Tables 6 and 7). Personality disorders were present in 5% (Puschner et al. 2004), 51% (Solbakken et al. 2012), and 80% (Jørgensen et al. 2000) of the patients.

2.5.2.3 Therapy and therapist characteristics

The short-term therapies were usually intended to last for 12-20 sessions, with individual, usually weekly, sessions lasting 45-60 minutes (Table 6). The median

Table 7. Characteristics of the short-term and long-term psychotherapy datasets (N=41) used in the evaluation of psychotherapy outcome prediction by suitability measures.

	Short-term therapy (N = 38)				Long-term therapy (N = 3)			
	N/Median		(%/Range)		N/Median		(%/Range)	
Country								
USA	17		(45)		-		-	
Canada	8		(21)		-		-	
Europe	11		(29)		3		(100)	
Netherlands		6		(16)		-		-
Norway		3		(8)		1		(33)
UK		1		(3)		-		-
Finland		1		(3)		-		-
Denmark		-		-		1		(33)
Germany		-		-		1		(33)
Australia & New Zealand	2		(5)		-		-	
Study design & therapy forms[1]								
Cohort study	24		(63)		3		(100)	
PP(I)		8		(21)		3		(100)
CT		6		(16)		-		-
BT		3		(8)		-		-
CBT		5		(13)		1		(33)
IPT		1		(3)		-		-
ET		1		(3)		-		-
Trial	14		(37)		-		-	
Therapy vs other		8		(21)				
PP(I)		2		(5)				
PP(S)		1		(3)				
CT		2		(5)				
CBT		1		(3)				
IPT		2		(5)				
Therapy vs therapy		6		(16)				
PP(I) vs PP(S)		1		(3)				
CBT vs IPT		3		(8)				
CBT vs ACT		1		(3)				
CBT vs BT		1		(3)				
Patient								
N (reported)[2]	52		(17-253)		153		(20-NR[3])	
N (analysed)[4]	50		(17-253)		153		(16-NR[3])	
Sex (% male)[5]	29		(0-53)		35		(29-45)	
Age (mean)[6]	36		(28-49)		35		(30-44)	
Diagnosis (%)[7]								
AXIS I	100		(65-100)		NR		(NR-NR)	
DEP		100		(12-100)		49		(47-50)
ANX		43		(4-100)		35		
AXIS II	54		(0-72)		51		(5-80)	
Therapy								
Duration (sessions)[8]	19		(4-34)		71		(69-72)	
Follow-up (months)[9]	12		(1-60)		24			
Therapist								
Training[10]								
T	26				3			
ST	12				1			
NT	1				-			
NR	5				-			

	Short-term therapy (N = 38)		Long-term therapy (N = 3)	
	N/Median	(%/Range)	N/Median	(%/Range)
Suitability measure[11]				
D1: P1	5		-	
D2: P2	4		1	
P3	4		-	
P4	1		-	
D3: P5	1		-	
P6	5		-	
P7	14		-	
D4: P8	5		-	
D5: P9	11		-	
P10	4		-	
P11	11		-	
P12	16		2	
D6: P13	5		1	
Symptom measure[12]				
V1	20		-	
V2	10		-	
V3	14		3	
Statistical method				
Correlation	27		1	
No adjustment		8		1
Baseline adjustment[13]		5		-
Other adjustment		5		-
Baseline + other		9		-
Modeling	15		2	
No adjustment		4		-
Baseline adjustment[13]		3		-
Other adjustment		-		2
Baseline + other		8		-

[1] ACT = Acceptance and Commitment Therapy, BT = Behavioral Therapy, CBT = Cognitive-Behavioral Therapy, CT=Cognitive Therapy, ET = Experiential Therapy, IPT = Interpersonal Therapy, PP (I) = Psychodynamic Psychotherapy (Interpretative), PP (S) = Psychodynamic Psychotherapy (Supportive)

[2] Size of the psychotherapy group for which the patient characteristics are reported. Not reported for 3 short-term and 1 long-term psychotherapy datasets. For 11 trials, patient characteristics are reported for study populationcomprising all the treatment groups.

[3] In the study by Puschner et al. 2004, 397 patients started PP(I) and 248 patients CBT but about 80% of them fulfilled the criteria for inclusion in the study. The sample size used in reporting patient characteristics or analysis was not reported.

[4] Size of the psychotherapy group used in the analysis. Not reported for 2 short-term and 1 long-term psychotherapy datasets.

[5] Based on 31/38 short-term psychotherapy datasets and 3/3 long-term psychotherapy datasets.

[6] Based on 35/38 short-term psychotherapy datasets and 3/3 long-term psychotherapy datasets.

[7] Axis I, Depressive disorders (DEP), Anxiety disorders (ANX), and Axis II based on 34/38, 26/38, 20/38, and 13/38 short-term psychotherapy datasets and 0/3, 2/3, 1/3, and 3/3 long-term psychotherapy datasets, respectively.

[8] Based on 21/38 short-term and 2/3 long-term psychotherapy datasets. Other datasets reported intended duration or nothing.

[9] Based on 10/38 short-term and1/3 long-term psychotherapy datasets involving follow-up.

[10] NT = Untrained, ST= Semi-trained, T = Trained.

[11] *D1. Nature of problems:* P1. Focus;
D2. Ego strength: P2. Modulation of affects, P3. Flexibility of interaction,P4. Self-concept in relation to ego ideal;
D3. Self-observing capacity: P5. Trial interpretation, P6. Reflective ability, P7. Motivation;
D4. Intelligence: P8. Intelligence;
D5. Intrapsychic and interpersonal behavior: P9. Coping styles and cognitive skills, P10. Defense styles, P11. Personality traits, P12. Interpersonal relationships;
D6. Overall suitability: P13. Overall suitability

[12] V1. Depressive symptoms, V2. Anxiety symptoms, V3. Global symptoms and functioning

[13] Adjustment for baseline symptom level

Patient suitability for short-term
and long-term psychotherapy

mean duration of therapy in studies reporting it (55%) was 19 sessions (range of mean durations 4-34) (Table 7), with the highest reported number of therapy sessions provided being 65 (Myhr et al. 2007). In one fourth of the datasets patients were followed up after the end of the therapy, with the follow-up varying from one month to five years, the median, and most typical, follow-up period being one year (Tables 6 and 7). Two studies with the longest follow-ups (4-5 years) only reported follow-up results (Husby et al. 1985, Husby 1985a; Høglend et al. 1992a, 1993, 1994, Høglend 1993). In all three studies on long-term therapies, the mean duration of therapy was around 70 sessions, although the range of therapy sessions attended varied greatly (Tables 6 and 7). One long-term therapy study (Solbakken et al. 2012) involved a 2-year post-treatment follow-up. Potential additional treatment received by the patients during the study protocol treatment or follow-up (e.g., psychotherapy, medication, and hospitalization) was seldom reported and never taken into account in the analyses.

In the case of 20 short-term psychotherapy datasets the therapies were delivered by trained and experienced therapists, in the case of five datasets by both trained and semi-trained therapists, who were currently under training or supervision, in the case of seven datasets by only semi-trained therapists, and in the case of one dataset by both trained and untrained therapists; five datasets did not report the training or experience of the therapists. All three long-term psychotherapy datasets involved trained and experienced therapists, but one of them (Jørgensen et al. 2000) also involved some semi-trained therapists.

2.5.2.4 Suitability and symptom measures

The vast majority of the suitability studies focused on suitability variables measuring intrapsychic and interpersonal behavior (Tables 6 and 7, Supplementary tables 1A-F). Of the 38 short-term psychotherapy datasets, 16 explored the prediction by the quality of patient's interpersonal relationships on their symptom development during and after therapy. The role of different personality traits as well as coping styles and cognitive skills in the prediction of psychiatric symptoms were both evaluated in 11 datasets, and defense styles in 4 datasets. Of the remaining suitability measures, motivation was studied the most, in 14 datasets. Focus, modulation of affects, flexibility of interaction with the interviewer, reflective ability, and intelligence were each studied based on 4-5 short-term psychotherapy datasets, whereas the role played by the patient's response to trial interpretation and his or her self-concept in outcome prediction were only explored in one dataset. The role of overall suitability measures, covering more than one aspect of suitability, was examined based on 5 short-term psychotherapy datasets. In long-term psychotherapies, on the other hand, only the role of three suitability measures on symptom development was evaluated: interpersonal relationships, based on two datasets, and modulation of affects and overall suitability, based on one dataset.

As most of the short-term psychotherapy datasets focused on patients suffering from depressive disorders, most of them (20 datasets) evaluated the development of patients' depressive symptoms (Table 7). Prediction by suitability measures of global symptoms and functioning was evaluated based on 14 datasets and on anxiety symptoms based on 10 datasets. The validated and widely used Beck Depression Inventory (BDI), Hamilton Depression Rating Scale (HDRS), and Symptom Checklist-90, Global Severity Index (SCL-90-GSI) were the most commonly used symptom measures: in 12, 10, and 11 short-term psychotherapy datasets, respectively (Table 6). All three long-term psychotherapy studies evaluated the prediction by suitability of global symptoms, two of them using the SCL-90-GSI (Tables 6 and 7).

2.5.2.5 Statistical methods

Most of the short-term psychotherapy datasets (71%) reported results in the form of correlation coefficients (Table 7, Supplementary tables 1A-F). Of these 27 datasets, 8 only reported simple correlations with no adjustment for the baseline symptom level or other confounding factors, 5 reported correlations adjusted for the baseline symptom level, 5 reported correlations adjusted for other confounding factors, and 9 reported correlations adjusted for both the baseline symptom level and other confounding factors. Four studies within the latter group also applied other modeling techniques. The remaining 11 datasets applied various modeling techniques to exploring the relationship between the suitability measure and symptom outcome, four with no adjustment, three adjusting for the baseline symptom level, and four adjusting for both the baseline symptom level and other confounding factors. Of the three long-term therapy studies, one reported results based on simple correlations with no adjustment and two based on modeling with adjustment for confounding.

2.5.3 Suitability measures as predictors of the outcome of short-term psychotherapies

2.5.3.1 Nature of problems

Evaluation of the nature of the patient's problems, their origin and scope, is necessary for the identification and establishment of the focus of psychotherapy treatment (see Section 2.2, Table 1).

Focus

The focality of the patient's chief complaint and the psychodynamic explanation given for it are considered to have an influence on suitability for different psychotherapies. Due to the limited time available in short-term therapies, well-circumscribed triangular problems stemming mainly from the oedipal phase are usually considered treatable with such therapies, whereas poorly circumscribed dyadic problems from the pre-oedipal phase, which are likely to anticipate greater

therapeutic difficulties such as dependency problems, are thought to require longer treatment (Sifneos 1972, Husby et al. 1985, Husby 1985a, Barth et al. 1988a, b). The relationship between the establishment of focus and the reduction in psychiatric symptoms was studied based on 5 datasets: 4 small to middle-sized (N = 18-43) studies on interpretative PP (Brodaty et al. 1982; Husby et al. 1985, Husby 1985a; Høglend 1993; Sigal et al. 1999) and one large (N =113) study on CBT (Myhr et al. 2007) (Supplementary table 1A). In all of these, the focus was observer-rated, either by the interviewer or the therapist, as part of the psychotherapy suitability assessment based on the suitability assessment scales described in Tables 3 and 4. In four datasets (Brodaty et al. 1982; Husby et al. 1985, Husby 1985a; Høglend 1993; Sigal et al. 1999) the observer evaluated the focal theme underlying the patient's psychological conflicts and the psychodynamic explanation for them based on the patient's description and prioritization of his or her key difficulties, whereas in one dataset (Myhr et al. 2007) the observer rated the patient's ability to focus in a session in a task-oriented fashion. No evidence of an association between the focus rating (varying from 2-point to 9-point) and improvement in general symptoms and functioning, either at the end of the therapy or at the follow-up, was found in any of the studies (Supplementary table 1A).

2.5.3.2 Ego strength

An individual's modulation of affects, flexibility of interaction, and sense of self are considered to reflect his or her ego strength, i.e., capacity to cope with competing, internal and external, demands and reality (see Section 2.2, Table 1).

Modulation of affects

Modulation of affects is defined as the individual's willingness and ability to access, express, regulate, differentiate, and integrate both positive and negative feelings. Good modulation of affects is considered necessary in order to be able to cope with and tolerate the distressing material, anxiety, and other negative affects introduced early on in time-limited, interpretative short-term therapies; patients with poor modulation of affects are likely to need longer and more supportive treatments to assist their recovery (Balint et al. 1972, Sifneos 1972, Davanloo 1978). Prediction by a patient's ability for modulation of affects on psychiatric symptoms was investigated based on 4 datasets (Brodaty et al. 1982; Sigal et al. 1999; Myhr et al. 2007; Cromer and Hilsenroth 2010) (Supplementary table 1B), all using the interview-based suitability assessment scales described in Tables 3 and 4.

The three studies on interpretative PP (Brodaty et al. 1982, Sigal et al. 1999, Cromer and Hilsenroth 2010) found no correlations between the patients' initial ability in terms of the modulation of affects and their general symptoms and functioning after the end of therapy (Supplementary table 1B). Cromer and Hilsenroth (2010) used the validated Capacity for Dynamic Process Scale (CDPS) (Tables 3 and 4) to evaluate the modulation of affects thoroughly from three

different aspects (perception, integration, and differentiation of affects) in a relatively large patient sample (N = 71) but none of them nor the overall affect factor was associated with changes in symptoms and functioning. Myhr et al. (2007), on the other hand, did find a single item of the validated Suitability for Short-term Cognitive Therapy (SSCT) Rating Scale (Tables 3 and 4), awareness and differentiation of emotions, to be correlated with change in global symptoms (r = 0.26) in CBT (Supplementary table 1B). This study included the greatest number of patients attending a certain therapy (N = 113) and a global symptom measure covering both general and diagnostic-specific symptoms. Furthermore, in this study anxious patients constituted the majority, whereas in the above-mentioned studies depressive patients constituted the majority. The small study by Brodaty et al. (1982; N = 18), which failed to find a correlation between depressive patients' ability to emote and change in psychiatric symptoms after 8-session PP, did find a strong correlation (r = 0.63) one year after the end of PP (Supplementary table 1B).

Flexibility of interaction

Flexibility of interaction is defined as the individual's ability to establish an adequate contact and collaboration with the interviewer. Flexible, active interaction with the interviewer is considered to reflect the patient's ability to form relationships, and thus to anticipate the formation of a good working alliance with the therapist within a short-term time frame, considered essential to the success of short-term interpretative therapies; patients who are more passive and submissive in their contact with the interviewer are likely to require more time for the formation of alliance, and longer and more supportive therapies in general (Sifneos 1972, Malan 1976a). The association between the flexibility of interaction and psychiatric symptoms was studied based on 4 separate datasets (Husby et al. 1985, Husby 1985a; Høglend 1993; Myhr et al. 2007; Cromer and Hilsenroth 2010) (Supplementary table 1B), all using interview-based suitability assessment scales (see Tables 3 and 4).

Myhr et al. (2007) found that an SSCT item, measuring quality of the patient-interviewer interaction and thus alliance potential, correlated with change in the general symptoms in CBT, whereas Cromer and Hilsenroth (2010) found no association between a similar CDPS item, evaluating therapeutical collaboration, and general symptoms in PP. The two other studies (Høglend 1993; Husby et al. 1985, Husby 1985a) on interpretative PP only evaluated the symptom change at the post-treatment follow-up: whereas Høglend found no association between the patient's involvement, a factor including both emotional and intellectual flexibility (Høglend et al. 1992a), and symptoms at the 2- or 4-year follow-up, Husby et al. found strong correlations between adequate contact with the interviewer and symptom relief both 2 and 5 years after the therapy had ended (r = 0.42 and r = 0.61, respectively).

Self-concept in relation to ego ideal

Self-concept is defined as the coherence and stability of the individual's representation and experiences of the self. Patients with a neurotic character structure and possessing a relatively mature, coherent, integrated and stable, self-concept are considered to benefit from short-term interpretative therapy, whereas patients with a less developed self-concept, underlying borderline or narcissistic character structures, are likely to require longer treatment in order to recover (Malan 1963, Horowitz et al. 1984). Only one study (Horowitz et al. 1984) explored the relationship between the development level of the patient's self-concept, evaluated with an observer-rated validated Self-Concept Rating Scale (Horowitz 1979, Horowitz et al. 1984), and change in psychiatric symptoms in PP, but did not find them to be significantly associated after adjustment for the symptom level at baseline ($r = 0.17$) (Supplementary table 1B).

2.5.3.3 Self-observing capacity

A person's capacity and willingness for self-observation can be assessed by evaluating his or her response to trial interpretation, reflective ability, and motivation (see Section 2.2, Table 1).

Response to trial interpretation

The individual's ability to respond to a trial interpretation given by an interviewer is determined based on his or her receptiveness to and elaboration of the interpretation (Davanloo 1980). The type and depth of interpretations depend upon the clinical situation, and a trial interpretation may therefore not be presented in every evaluation interview; focus interpretation, i.e., interpretation of the suggested main problem, is often among the first interpretations in relation to which a patient's response, a recognition and elaboration of the problem area, can be evaluated. The ability to accept and elaborate on interpretations is naturally considered essential to attending any short-term interpretative therapy, whereas more supportive and longer treatments are likely to be beneficial for patients unresponsive to trial interpretation. In the only published study (Brodaty et al. 1982) exploring the relationship between the patient's response to trial interpretation and changes in symptoms, no correlation was found between the patient's ability to accept interpretations and the change in general psychiatric symptoms, neither at the end of PP nor at the 1-year follow-up (Supplementary table 1C). This study was, however, very small (N = 18), and the interview-based suitability assessment scale applied in it unvalidated (see Table 4).

Reflective ability

Reflective ability encompasses a person's psychological mindedness and insight, i.e., his or her capacity for introspection, and the ability to elaborate and analyze one's inner psychic life, thoughts, feelings, and actions, to understand how these are related, both to one another and to past and current experiences, and to communicate

them verbally; in therapy, the degree of this communication via consideration of the available time span is also important (Appelbaum 1972, Davanloo 1980). It is hypothesized that a good reflective ability enhances the patient's engagement with therapy and the tolerance of anxiety and stress occurring during the therapy, therefore increasing his or her ability to benefit from short-term, especially insight-oriented, psychotherapy, whereas lack of reflective ability is likely to indicate the need for longer and more supportive treatment (Conte et al. 1990, Shill and Lumley 2002). The relationship between the level of reflective ability and psychiatric symptoms was investigated based on 5 separate datasets (Supplementary table 1C); three of them (Brodaty et al. 1982; Myhr et al. 2007; Cromer and Hilsenroth 2010) applied the previously described interview-based suitability assessment scales (Tables 3 and 4), one validated observer-rated method (Piper et al. 1998), and one validated self-report scale (Kronström et al. 2009) in the assessment of reflective ability.

Two small studies on interpretative PP (Brodaty et al. 1982; N = 18, Kronström et al. 2009; N = 19) found no association between the patient's psychological mindedness and change in symptoms by the end of the therapy. Kronström et al. (2009) did not find an association at the 8-month post-treatment follow-up either, whereas Brodaty et al. (1982) found a significant association at the 12-month follow-up (r = 0.45). Further analysis by Brodaty et al. (1982) showed that this correlation was stronger in men (r = 0.74), and the percentage of men in this study was higher than in the study by Kronström et al. (2009). However, Brodaty et al. (1982) evaluated psychological mindedness based on an interview and in relation to general psychiatric symptoms, whereas Kronström et al. (2009) used the 45-item self-report Psychological Mindedness Scale (Conte et al. 1990) and evaluated depressive symptoms. Piper et al. (1998) found psychological mindedness, rated by an observer via the Psychological Mindedness Assessment Procedure (McCallum and Piper 1997), to correlate with general psychiatric symptoms (r = 0.26) in interpretative PP (N = 64) but not in supportive PP (N = 66). Psychological mindedness did not, however, correlate with depressive or anxiety symptoms in either treatment group. A relatively large (N = 71) study by Cromer and Hilsenroth (2010) evaluated three aspects of reflective ability via the CDPS, appearing introspective, manifesting insight, and manifesting verbal fluency, and found that patients who manifest more insight prior to the therapy experienced greater improvement in their general symptoms (r = 0.26) and functioning (r = 0.27) in PP. The insight factor formed based on these three single items was also statistically significantly correlated with functioning (r = 0.24) but not with symptoms (r = 0.17). The largest study (N = 113) by Myhr et al. (2007) found no association between accessibility of automatic thoughts, i.e., negative, self-critical thinking relevant to the therapy focus, and the change in general symptoms in CBT.

Motivation

Motivation as a psychotherapy suitability selection criterion is a multi-faceted concept that covers motivation and readiness for change, motivation for psychotherapy as a treatment option, own interest in and commitment to therapy, willingness to make reasonable sacrifices in terms of time and money, recognition of the psychological etiology of problems and agreement with the treatment rationale, motivation for introspection and self-understanding, motivation for active and open collaboration, the patient's perseverance, and realistic expectations of the therapy outcome (Sifneos 1972, Malan 1976a, Davanloo 1978, 1980). All of these aspects of motivation have been considered relevant to the successful outcome of both short-term and long-term psychotherapy; excellent motivation is, however, particularly emphasized in relation to achieving a positive outcome in time-limited short-term psychotherapy. The association between different aspects of motivation and psychiatric symptoms was studied based on 14 datasets; in 6 of them using interview-based evaluation (Brodaty et al. 1982; Horowitz et al. 1984; Husby et al. 1985, Husby 1985a; Høglend 1993; Sigal et al. 1999; Myhr et al. 2007) and in 8 using patient self-report questionnaires (Sotsky et al. 1991; Keijsers et al. 1994a; Keijsers et al. 1994b; Zuroff et al. 2007; Kampman et al. 2008; McBride et al. 2010; Steketee et al. 2011; Lewis et al. 2012) (Supplementary table 1C).

Five studies, three on interpretative PP (Brodaty et al. 1982, Horowitz et al. 1984, Sigal et al. 1999), one study on CBT (Kampman et al. 2008) and one study on CBT and IPT (Sotsky et al. 1991), found no association between different aspects of patient's pre-treatment motivation and post-treatment psychiatric symptoms. These were small to medium-sized studies (N = 18-61), except for the relatively large study by Kampman et al. (2008; N = 129). In the two studies by Keijsers et al. (1994a, b; N = 40-60) on the treatment of panic disorder and obsessive-compulsive disorder (OCD) with behavioral therapy (BT), motivation – measured using the reliable Nijmegen Motivation List (NML; Keijsers et al. 1991, 1999) also used by Kampman et al. (2008) – emerged as a statistically significant predictor of remission from anxiety symptoms and reduction of obsessive fear, but not of compulsive behavior. Keijsers et al. (1994a, b), however, evaluated motivation only using the first subscale of NML, measuring the patient's willingness to participate, whereas Kampman et al. (2008) also used the two other subscales, measuring demoralization and a reserved attitude towards treatment. In the larger study on panic disorder (Keijsers et al. 1994a) the findings also remained significant after adjustment for potential confounding factors, whereas in the smaller study on OCD (Keijsers et al. 1994b) they did not. In the study by Zuroff et al. (2007), autonomous motivation, representing the patient's own intrinsic motivation for therapy, but not controlled motivation, representing both external and introjective reasons for participating in therapy, predicted lower post-treatment depression severity in both IPT and CBT. This finding was confirmed in a later study by McBride et al. (2010), which used the same Autonomous and Controlled Motivations for Treatment Questionnaire,

adapted from the validated Treatment Self-Regulation Questionnaire (Williams et al. 1998), and also found that autonomous motivation, but not controlled motivation, predicted overall remission from baseline-adjusted depressive symptoms in IPT. Further subgroup analyses by McBride et al. (2010), however, showed that higher autonomous motivation predicted remission among those with less recurrent depression but not among those with highly recurrent depression, and that controlled motivation was a significant negative predictor of remission in both subgroups. In a large study (N = 113) by Myhr et al. (2007), acceptance of personal responsibility for change was statistically significantly correlated (r = 0.21) to a change in general symptoms in CBT; correlations for the patient's optimism/pessimism regarding therapy (r = 0.19) and compatibility with cognitive rationale (r = 0.17) were in the same range but did not reach statistical significance. In the largest study (N = 173) by Lewis et al. (2012) low initial motivation for change, evaluated using the validated Stages of Change Schedule (McConnaughy et al. 1983), also known as the University of Rhode Island Change Assessment Questionnaire (URICA; McConnaughy et al. 1989, Greenstein et al. 1999), predicted a reduction in baseline-adjusted depressive symptoms in CT. In another study by Steketee et al. (2011), on the other hand, a high initial motivation for change, evaluated using the URICA, was found to predict a reduction in baseline-adjusted OCD symptoms at the end of CBT, but no longer at the 1-year post-treatment follow-up; in this study, however, the motivation for change was only evaluated after session 4 or 6 and may thus have already changed. Regarding other post-treatment follow-up findings, no association was identified between motivation and general symptoms at the 1-year follow-up by Brodaty et al. (1982), at the 2-year follow-up by Høglend (1993) or Husby et al. (1985) or at the 4-year follow-up by Høglend (1993); a significant correlation (r = 0.34) was, however, found at the 5-year follow-up by Husby (1985a).

2.5.3.4 Intelligence

Intelligence encompasses a person's complex mental abilities in acquiring and applying knowledge, in abstract reasoning, and in analyzing and solving problems (see Section 2.2, Table 1).

Intelligence

Intelligence can be defined in many different ways, but is traditionally defined and measured using Intelligence Quotient (IQ) tests, such as the Stanford-Binet (Terman 1916) and the Wechsler Adult Intelligence Scale (WAIS; Wechsler 1955). Whereas in therapies with a cognitive focus the IQ-based conception and measurement of intelligence remains common, within a psychodynamic frame a broader definition of intelligence based on person's problem-solving capacity is often adapted (Davanloo 1980, Barth et al. 1988b). Above-average or average intelligence is considered an important suitability criterion for short-term psychotherapy (Davanloo 1980). Patients with superior intelligence may, however, be resistant to psychotherapy. The

Patient suitability for short-term
and long-term psychotherapy

association between intelligence or problem-solving capacity and symptoms was explored based on 5 datasets: in three datasets on CT, CBT, or BT (Haaga et al. 1991; Rizvi et al. 2009; Fournier et al. 2009) based on validated self-report instruments, highly correlated with WAIS IQ score, and in two datasets on PP (Husby et al. 1985, Husby 1985a; Høglend et al. 1994) based on an interview evaluating the patient's problem-solving capacity (Supplementary table 1D).

Haaga et al. (1991) did not find that crystallized intelligence, i.e., depth and breadth of general knowledge, measured using the 40-item self-administered WAIS-Clarke vocabulary test (Paitich and Crawford 1970), or fluid intelligence, i.e., the capacity to think logically and solve novel problems irrespective of acquired knowledge, measured using the 40-item Abstractions subscale of the Shipley Institute of Living Scale, also known as the Shipley-Harford Living Scale (Shipley 1940), predicted remission from depressive symptoms among those with depressive disorders (N = 76) or remission from anxiety symptoms among those with anxiety disorders (N = 30) in CT. Fournier et al. (2009), however, found that lower intelligence, measured using the 20-item verbal (i.e., crystallized) intelligence and 10-item analytic intelligence subscales of the Shipley-Harford Living Scale (Shipley 1940), predicted a poorer response to CT in the treatment of patients suffering from depressive disorders. Rizvi et al. (2009) studied women with post-traumatic stress disorder (PTSD) and found that lower general intelligence, measured using the 50-item Quick Test (Ammons and Ammons 1962), was associated with a higher dropout rate but not with PTSD symptomatology either CBT or BT. Husby et al. found a statistically significant correlation between neurotic patients' problem-solving capacity and symptom relief both at the 2-year (r = 0.34; Husby et al. 1985) and the 5-year follow-up (r = 0.44; Husby 1985a). Høglend et al. (1994), on the other hand, did not find any overall association between patients' problem-solving capacity (including evaluation of self-understanding and insight) and general symptoms at the 2- or 4-year follow-up. Despite the relatively small sample size, they did, however, find an interaction between the patient's level of problem-solving capacity and number of therapy sessions (range 9-53), so that for patients with a low or moderate level of problem-solving capacity a longer treatment was significantly correlated with greater symptom reduction at a 4-year follow-up, in line with theoretical suggestions (Sifneos 1972).

2.5.3.5 Intrapsychic and interpersonal behavior

A person's coping styles and cognitive skills, defense styles, personality traits and interpersonal relationships are considered to represent his or her intrapsychic and interpersonal behavior (see Section 2.2, Table 1).

Coping styles and cognitive skills

Coping styles refer to a person's characteristic ways of thinking, feeling, and acting when confronted with problems, aimed at mastering the stress and anxiety that has

arisen and functioning better in the situation (Lazarus and Folkman 1984, Zeidner and Endler 1996). Cognitive skills refer to a person's ability to (correctly) perceive, process, reason, and relate new information, and the related habitual patterns of thinking, feeling, and behaving (Beck et al. 1979). The level of coping styles and cognitive skills is often evaluated on a continuum from adaptive and functional to maladaptive and dysfunctional, using measures of ways of coping, feelings of control, learned resourcefulness, self-efficacy, attributional style, hopelessness, and dysfunctional attitudes. Adaptive and functional coping styles and cognitive skills are considered beneficial to achieving a positive outcome in any psychotherapy, but in particular in short-term therapies, especially those of cognitive orientation. The more maladaptive and dysfunctional the patient's coping styles and cognitive skills the longer the required treatment likely needs to be, in order to properly address the deficits in the patient's ability to cope and explore new coping styles, and to identify distorted thinking and beliefs, in order to modify them and one's behavior and way of relating to others. The association between coping styles and cognitive skills and psychiatric symptoms was explored in 11 separate datasets, all on CT or CBT and/or IPT and all using self-report instruments (Zuckerman et al. 1980; Simons et al. 1985; Jarrett et al. 1991a, b; Sotsky et al. 1991, Blatt et al. 1995; Simons et al. 1995; Hardy et al. 2001; Clark et al. 2003b; Spangler et al. 2007; Fournier et al. 2009; Carter et al. 2011; Lewis et al. 2012) (Supplementary table 1E).

The association between dysfunctional attitudes, self-reported using the validated Dysfunctional Attitudes Scale (DAS; Weissman 1979), and depressive symptoms was studied in all except one (Zuckerman et al. 1980) datasets, with somewhat contradictory findings. Two large studies (Clark et al. 2003b, Lewis et al. 2012; N = 100-173) did not find dysfunctional attitudes to be associated with depressive symptoms in CT or CBT. A small study by Simons et al. (1985; N = 19) did not find that dysfunctional attitudes, or dysfunctional thinking or cognitive processing, predicted symptom development in CT either. In a later, larger study (N = 53), Simons et al. (1995) did find higher levels of dysfunctional attitudes to be associated with a poorer response to CBT; this association, however, only appeared among those patients who had not experienced a severe negative life event. Jarrett et al. (1991a; N = 37) also found higher DAS scores to predict poorer response to CT; further analysis showed that single patients with high DAS scores responded least. Hardy et al. (2001; N = 24) did not find the three subscales of DAS (achievement, dependency, self-control) to correlate with depressive symptoms in CT. According to Spangler et al. (1997; N = 53) DAS achievement and interpersonal subscales were not associated with depressive symptoms in CBT either. Sotsky et al. (1991) found patients with lower initial cognitive dysfunction, evaluated based on DAS, to be more responsive to CBT, whereas the level of cognitive dysfunction did not seem to play a role in IPT. Based on the same dataset, Blatt et al. (1995) found that perfectionism but not a need for approval, both factors based on DAS, predicted higher baseline-adjusted depressive and global symptoms both in CBT and IPT.

Fournier et al. (2009), on the other hand, did not find that perfectionism or the need for approval predicted depressive symptoms in CBT, when adjusted for the baseline symptom level and for each other. Finally, according to Carter et al. (2011) dysfunctional attitudes predicted depressive symptoms in both CBT and IPT.

Zuckerman et al. (1980) did not find Frank's Mastery Scale, adapted from Seeman's (1967) Powerlessness Scale and measuring the patient's self-reported assertive behaviors and feelings of control over various aspects of life, to be significantly related to depressive symptoms upon treatment termination or at the 1-year post-treatment follow-up of IPT. A closely related concept of learned resourcefulness, measured using the validated Self-Control Schedule (Rosenbaum 1980) covering four categories of self-control skills (cognitive strategies, problem solving, delay of gratification, and self-efficacy expectations), emerged as a strong predictor of depressive symptoms (r = 0.53) in CT in a small study by Simons et al. (1985; N = 19), but not in a larger study by Jarrett et al. (1991b; N = 37), although the same validated outcome measure (BDI), adjusted for the baseline level, was used in both studies. Self-efficacy, measured with the validated Self-Efficacy Scale (Sherer et al. 1982), and self-esteem, measured with the validated Rosenberg Self-esteem Scale (Rosenberg 1965), were not associated with depressive symptoms in two relatively large studies on CT and CBT (Clark et al. 2003b, Fournier et al. 2009; N = 60-100). Neither were any associations found between the patient's attributional style (i.e., tendency to attribute negative and positive events to internal versus external, stable versus unstable, and global versus specific causes), measured using the validated Attributional Style Questionnaire (ASQ; Seligman et al. 1979, Peterson et al. 1982), and depressive symptoms in three studies on CT or CBT (Jarrett et al. 1991a, Clark et al. 2003b, Fournier et al. 2009). However, a study by Spangler et al. (1997), using a modified version of the ASQ allowing the evaluation of interpersonal and achievement attributional styles, found that a global-stable interpersonal attributional style in combination with a negative interpersonal life event predicted worse response to CBT. On the other hand, a closely related concept of hopelessness, self-reported using the validated Hopelessness Scale (Beck et al. 1974), was not found to relate to depressive symptoms in CT or CBT (Simons et al. 1985, Fournier et al. 2009, Lewis et al. 2012).

Defense styles

Defense styles are defined on the basis of the availability and integration of individual regulating functions, defense mechanisms, which are aimed at alleviating anxiety and maintaining mental balance (American Psychiatric Association 1994, OPD Task Force 2001). Individual defense mechanisms are typically grouped under three main defense styles: mature, neurotic, and immature (Andrews et al. 1993, Abraham et al. 2001). Patients with an adaptive, flexible and well-integrated, mature defense style are considered suitable candidates for short-term interpretative, anxiety-provoking psychotherapies, whereas patients with maladaptive, rigid and

less integrated, immature or neurotic defense styles are likely to need supportive and/or longer treatment (Davanloo 1980, Van et al. 2009). The relationship between patients' defense styles and psychiatric symptoms was evaluated in 4 studies based on separate datasets (Hersoug et al. 2002, Myhr et al. 2007, Kronström et al. 2009, Van et al. 2009) (Supplementary table 1E), both based on observer-rated and self-reported methods.

The two smallest studies on interpretative PP by Kronström et al. (2009; N = 19) and Hersoug et al. (2002; N = 39) did not find any association between self-reported, with validated 88-item Defense Style Questionnaire (DSQ; Bond et al. 1983, 1989, Andrews et al. 1989, 1993), or observer-rated, with validated Defense Mechanisms Rating Scale (Perry 1990), mature, neurotic, immature, or overall defensive functioning and depressive symptoms or global symptoms and functioning neither at the end of PP or at 8-month follow-up. In a larger study by Van et al. (2009; N = 69-81) both self-reported, with a 42-item DSQ (Trijsburg et al. 2000), and observer-rated, with the validated Developmental Profile (Abraham et al. 2001), overall defensive functioning were found to predict the remission from depressive symptoms in supportive PP; of the self-reported defense styles, adaptive mature defense style, and of the observer-rated defense styles, a lower maladaptive symbiotic level (giving up, apathetic withdrawal) and a higher maladaptive level of rivalry (repression, affect denial) predicted remission. Myhr et al. (2007) found a statistically significant correlation (r = 0.25) between the level of a patient's security operations, i.e., the intensity of the defensive information processing strategy aimed at blocking anxiety-provoking material, and the change in general psychiatric symptoms in CBT.

Personality traits

Personality traits cover both global personality constructs as well as specific personality types and traits. Adaptive and flexible personality dimensions have been suggested as indicators for suitability for psychotherapy, time-limited short-term psychotherapy in particular, whereas patients possessing nonadaptive personality dimensions are likely to require longer treatment in order to recover. The relationship between personality traits and psychiatric symptoms was investigated based on 11 separate datasets (Zuckerman et al. 1980; Brodaty et al. 1982; Sigal et al. 1999; Clark et al. 2003a, b; Blom et al. 2007b; Marshall et al. 2008; Joyce et al. 2007; Rizvi et al. 2009; Kronström et al. 2011; Cromer and Hilsenroth 2010; Wolitzky-Taylor et al. 2012) (Supplementary table 1E), mainly based on self-report instruments.

Clark et al. (2003a, b; N = 100-108) studied the association between 15 dimensions of personality functioning, self-reported using a validated 375-item Schedule for Nonadaptive and Adaptive Personality (Clark 1993), and depressive symptoms in CT: although 6 of them correlated with pre-treatment depression severity, none of them correlated with post-treatment depression severity in CT. As

personality assessment is likely to tap both state and trait variance and as their further studies demonstrated that changes in depression severity correlated with changes in personality states but not with stable personality traits, Clark et al. (2003a) came to the conclusion that the state variance was masking the trait variance, and therefore recommended that personality traits be separated from personality states in the prediction of depression. Joyce et al. (2007) studied another validated self-report measure, the Temperament and Character Inventory (TCI; Cloninger et al. 1994), which assesses people on the basis of four dimensions of temperament (novelty seeking, harm avoidance, reward dependency, and persistency) and three dimensions of character (self-directedness, self-transcendence, and cooperativeness). Only low persistence was found to predict a worse treatment outcome in CBT (N = 80), whereas in IPT (N = 87) high harm avoidance and low self-directness in particular, as well as low novelty seeking and reward dependency predicted a worse outcome. A small study by Kronström et al. (2011; N = 19) examined the potential role of TCI in interpretative PP but found no statistically significant associations between TCI personality dimensions and post-treatment depression scores. Marshall et al. (2008) studied yet another self-report questionnaire, a validated 66-item Depressive Experiences Questionnaire (Blatt et al. 1976), which assesses two characterological configurations presumably associated with depression: self-criticism and dependency. They found self-criticism to be a significant predictor of a higher level of post-treatment depressive symptoms in IPT (N = 35) but not in CBT (N = 37), and dependency to be a nearly significant predictor of a worse treatment outcome in CBT but not in IPT. Blom et al. (2007b) studied five personality dimensions (neuroticism, extraversion, openness, agreeableness, and conscientiousness) of the Big Five model, using NEO-Five Factor Inventory, a 60-item short form (Costa and McCrae 1992) of the validated self-reported 240-item NEO Personality Inventory – Revised (NEO-PI-R; Costa and McCrae 1992), but found none of these to be predictive of depressive symptoms in IPT. Zuckerman et al. (1980) did not find neuroticism or extraversion, measured using validated self-reported Maudsley Personality Inventory (Eysenck 1959), nor feelings of guilt, evaluated using the guilt questions derived from the validated self-reported Buss-Durkee Guilt-Hostility Scale (Buss and Durkee 1957), to predict depressive symptoms in IPT either, neither upon treatment termination nor at the 1-year post-treatment follow-up. Wolitzky-Taylor et al. (2012) studied the association between the neurotic personality trait, using the NEO-PI-R, and anxiety symptoms in CBT and Acceptance and Commitment Therapy (ACT) both at the end of treatment and at 6- and 12-month post-treatment follow-ups, and found higher baseline neuroticism to predict higher levels of baseline-adjusted anxiety across treatment groups and assessment time points. Of the other personality traits, anger-hostility, evaluated using validated self-reported State-Trait Anger Expression Inventory (Spielberger and Sydeman 1994), was not found to be associated with PTSD symptoms in either CBT or BT (Rizvi et al. 2009). Of the interview-based

evaluations, Brodaty et al. (1982) studied the role of patient attractiveness and likability, and Cromer and Hilsenroth (2010) studied a similar variable evaluating whether a patient was offering a positive relationship, but neither of them was found to predict general symptoms and functioning in interpretative PP. The similar cooperativeness dimension of self-report TCI was not found to correlate with depressive symptoms in CBT, IPT, or PP either (Joyce et al. 2007, Kronström et al. 2011). Finally, Sigal et al. (1999) studied the relationship between the creativity factor and improvement in symptoms and functioning in PP, but found none.

Interpersonal relationships

The quality of interpersonal relationships is evaluated based on a person's developmental level of relational patterns, current and past capacity for interpersonal functioning, and specific interpersonal behavioral traits. Regarding psychotherapy suitability, a history of meaningful, altruistic, adaptive give-and-take relationships is particularly emphasized, and at least one such relationship with a significant other is considered a requisite for short-term psychotherapy, indicating the patient's ability and willingness to invest in therapy and develop a confiding therapeutic relationship with the therapist (Sifneos 1972, Davanloo 1980). The association between the history and quality of a patient's interpersonal relationships and the development of his or her psychiatric symptoms during and after short-term psychotherapy was reported based on 16 separate datasets (Emmelkamp 1980; Husby et al. 1985, Husby 1985a; Sotsky et al. 1991; Høglend et al. 1992a, 1993, Høglend 1993; Paivio and Bahr 1998; Piper et al. 1998; Sigal et al. 1999; Hilliard et al. 2000; Hardy et al. 2001, Saatsi et al. 2007; Clark et al. 2003b; McBride et al. 2006; Carter et al. 2011; Myhr et al. 2007; Van et al. 2008; Cromer and Hilsenroth 2010; Renner et al. 2012) (Supplementary table 1E). The nature of interpersonal relationships has been evaluated both using measures specifically designed for that purpose and as part of larger suitability assessment scales covering several aspects of a patient's psychological functioning (see Table 3).

The patient's history and quality of interpersonal relationships is often evaluated via their object relational functioning, reflecting a relatively enduring tendency to establish certain types of actual relationships with others and internal representations of relationships, according to the degree of their maturity (Azim et al. 1991). Piper et al. (1998) found a statistically significant association between patients' quality of object relations, evaluated on an overall dimension from primitive to mature based on a validated interview-based Quality of Object Relations Scale (QORS; Azim et al. 1991), and the reduction of both anxiety and depressive symptoms in interpretative PP ($r = 0.38$ and 0.26, respectively) but not in supportive PP ($r = 0.00$ and 0.07, respectively). On the other hand, no association between QORS and general psychiatric symptoms was found in either group. Van et al. (2008) studied 9 developmental levels of object relational functioning and a total object relational functioning (TORF) score based on them, using a validated interview-based

Developmental Profile (Abraham et al. 2001), and found higher scores on the adaptive levels of individuation and solidarity, lower scores on the disadaptive level of lack of structure, and more mature TORF to be associated with a better symptom outcome in supportive PP. In further multivariate modeling, only individuation remained as a significant predictor of outcome, however.

The three commonly used validated self-report questionnaires in the assessment of interpersonal and social functioning and interaction are the Structural Analysis of Social Behavior (SASB) INTREX questionnaire (Benjamin 1974, 1983), allowing assessment of the quality of any interpersonal relationship, Social Adjustment Scale (SAS-SR; Weissman and Bothwell 1976), assessing an individual's functioning in his or her major roles (work, social and leisure activities, partner, parent, family unit, extended family), and Inventory of Interpersonal Problems (IIP; Horowitz et al. 1988, 2000), assessing the extent of problematic behaviors, thoughts, and feelings in one's significant relationships. According to Hilliard et al. (2000), the quality of a patients' early parental relations, assessed using the SASB INTREX questionnaire, was associated with their symptom change in interpretative PP (r = 0.27). Clark et al. (2003b) studied both SAS-SR and IIP, but did not find any correlation between their total scores and depressive symptoms in CT in a sample of 100 patients. According to Sotsky et al. (1991) low social dysfunction, measured based on a social and leisure activities subscale of SAS-SR, predicted a reduction in depressive symptoms in IPT but not in CBT. Carter et al. (2011) studied the prediction of modified SAS-SR (Cooper et al. 1982), assessing interpersonal relations, role performance, the amount of friction with others, and inner feelings and satisfaction, on depressive symptoms in IPT and CBT treatment groups combined, but found that only inner feelings and satisfaction subscale showed a nearly significant prediction despite the large sample size (N = 177). Emmelkamp (1980) studied the influence of interpersonal problems, both problems with the significant partner, self-reported using the Marital Deprivation Scale, adapted from the validated Marital Attitude Evaluation Scale (Schutz 1973), and assertiveness, self-reported using the validated Adult Self-Expression Scale (Gay et al. 1975), on BT treatment for agoraphobics, and found assertiveness to predict a better response at the end of a 4-week period of treatment but no longer at a 4-week post-treatment follow-up in a small sample of 17 patients. Paivio and Bahr (1998) studied the circumplex model of IIP (IIP-C; Alden et al. 1990), i.e., the two main dimensions of interpersonal behavior (affiliation or communion, and dominance or agency) and the eight subscales (domineering, competitive, overly cold, socially avoidant, nonassertive, exploitable, overly nurturant, and intrusive), but found none of the scores for these interpersonal variables to be related to symptom outcome in experiential therapy (ET). Renner et al. (2012) studied a three-factor structure of IIP-C (general interpersonal stress factor, agency factor, and communion factor) and found elevated baseline general interpersonal distress factor scores to be significantly associated with higher depressive symptoms throughout CT; although patients with higher agency factor

scores had lower symptoms in the middle of CT and slightly lower symptoms at the end of CT, neither agency nor communion factor scores were statistically significantly related to the symptoms. Hardy et al. (2001) found patients' interpersonal style, developed from IIP, to correlate strongly with depressive symptoms in CT, also after adjustment for the baseline symptom level, in a relatively small sample of 24 patients: both patients with an underinvolved interpersonal style, i.e., highly avoidant of relationships, and patients with an overinvolved interpersonal style, i.e., influenced too much by their relationships, showed less symptom improvement during post-treatment (r = 0.62 and r = 0.51, respectively). On the other hand, McBride et al. (2006) found that patients scoring higher on attachment avoidance, measured based on a validated self-report Relationship Scales Questionnaire (Griffin and Bartholomew 1994), showed a significantly greater reduction in depressive symptoms in CBT than in IPT. Saatsi et al. (2007) further studied the role of a secure, avoidant, and ambivalent interpersonal style, based on IIP and Attachment Vignettes (Hazan and Shaver 1987), in psychotherapy outcome prediction based on the same dataset. They found that secure patients had significantly lower-level post-treatment depressive symptoms than patients with an avoidant and ambivalent interpersonal style; a secure interpersonal style was a significant outcome predictor, including after adjustment for baseline symptoms. However, further studies by Hardy et al. (2001) and Saatsi et al. (2007) demonstrated that the impact of an interpersonal style was mediated through therapeutic alliance.

Regarding the larger suitability assessment scales, including interpersonal behavior as one dimension in the evaluation, Myhr et al. (2007) found that the patient's alliance potential, evaluated based on the history of meaningful relationships as part of the SSCT Rating Scale (Table 3), correlated with an improvement in psychiatric symptoms in CBT. Sigal et al. (1999) found no association between the stability of a patient's relationships or the presence of transference manifestations, which forms part of the PP Checklist (Table 3), and psychiatric symptoms and functioning. Cromer and Hilsenroth (2010) studied the interview-based CDPS (Table 3) but did not find that the patient's ability to differentiate interpersonal events or the overall relational factor correlated with general psychiatric symptoms and functioning in PP. Husby et al. did not find a statistically significant correlation between the existence of at least one adequate give-and-take relationship, assessed on the Initial Evaluation Form for STAPP (Table 3), and symptom relief at the 2-year (Husby et al. 1985) or at the 5-year (Husby 1985a) follow-up after PP. Høglend (1993), on the other hand, did find a statistically significant correlation between the quality of interpersonal relations, forming part of a further modified version of Sifneos' selection criteria for STAPP (Table 3), and the change in symptoms and functioning at 2 years (r = 0.35), but no longer at 4 years (r = 0.17), after the end of PP. When baseline symptoms and

personality disorders were adjusted for, the association at the 2-year follow-up ($r = 0.30$) was no longer significant, however.

2.5.3.6 Overall suitability

Since patients may possess some of the capacities considered relevant to psychotherapy suitability but not others, and different capacities may balance each other, patients with good overall suitability are likely to benefit from psychotherapy, short-term psychotherapy in particular for which these capacities have been especially emphasized (see Section 2.2, Table 1). Suitability assessment scales (Table 3), which cover several dimensions of suitability, i.e., nature of problems, ego strength, self-observing capacity, intelligence, and intrapsychic and interpersonal behavior, can be used to evaluate overall suitability. The association between overall patient suitability and the development of psychiatric symptoms in short-term psychotherapy was investigated in 5 studies based on separate datasets, using three of the suitability assessment scales (Brodaty et al. 1982; Alpher et al. 1990; Safran et al. 1993; Myhr et al. 2007; Cromer and Hilsenroth 2010) (Supplementary table 1F). Alpher et al. (1990) found a higher, and thus better, total CDPS score to predict a larger change in psychiatric symptoms and functioning in interpretative PP ($r = 0.45$-0.56) in a sample of 25 patients; Cromer and Hilsenroth (2010) did not, however, find a significant association between the same overall suitability score and the same or similar symptom measures ($r = 0.20$-0.23) in a larger sample of 71 patients. Strong correlations between mean suitability based on the SSCT Rating Scale and depressive, anxiety, and general symptoms after CBT were found both by Safran et al. (1993) and Myhr et al. (2007), including after adjustment for the baseline symptom level and confounding factors. Brodaty et al. (1982) did not find that the total predictor variable (TPV), calculated by summing up the values of the six items of his selection criteria, correlated with the change in general psychiatric symptoms by the end of PP in a small sample of 18 patients; TPV did, however, correlate strongly ($r = 0.55$) with the change in psychiatric symptoms by the 1-year post-treatment follow-up, with patients with better suitability gaining more from the therapy.

2.5.4 Suitability measures as predictors of the outcome of long-term psychotherapies

The association between pre-therapy psychological suitability and the development of symptoms in long-term psychotherapy has been evaluated based on only three datasets in three separate studies (Jørgensen et al. 2000; Puschner et al. 2004; Solbakken et al. 2012). Of the psychological suitability measures (see Table 1), only the role of modulation of affects (Solbakken et al. 2012), interpersonal relationships (Jørgensen et al. 2000; Puschner et al. 2004), and overall suitability (Jørgensen et al. 2000) has been evaluated. Solbakken et al. (2012) found that patients with more severe impairment of affect consciousness and integration, evaluated using a semi-

structured validated Affect Consciousness Interview (Monsen et al. 2008), experienced greater improvement in general psychiatric symptoms in open-ended psychodynamically oriented therapy, lasting on average 72 sessions, in a large sample of 153 patients. Puschner et al. (2004) studied the prediction by interpersonal problems, affiliation and dominance scores of IIP, of the rate of symptom change in CBT (25-50 sessions) and PP (50-80 sessions), and found low affiliation, i.e., hostility, and interaction between affiliation and dominance, i.e., hostile submission, to predict the highest rate of symptom change in PP. Jørgensen et al. (2000) found significant correlations between both positive interpersonal contributions, referring to open, friendly, and confident interpersonal behavior, and negative interpersonal contributions, referring to controlling or withdrawn interpersonal behavior, evaluated using SASB, and a change in general psychiatric symptoms in PP with an average duration of 69 sessions (r = 0.54-0.57) in a small sample of 16 patients. They did not, however, find a significant association between the overall Dynamic Assessment Interview (DAI) rating (Table 3) and symptom change.

2.5.5 Summary of findings on the prediction by suitability measures of the psychotherapy outcome

2.5.5.1 Availability of the suitability studies

Several studies have evaluated the prediction by the patients' pre-treatment psychological suitability measures of the outcome of individual psychotherapies, which has allowed evaluation of how the proposed criteria for suitability for psychotherapy (Table 1) work in practice. With the exception of three studies, all suitability research has, however, focused on short-term psychotherapies.

All of the psychological suitability measures presented in Table 1 have been studied for their association with the development of psychiatric symptoms during individual short-term therapy. Some aspects of psychological suitability have been studied much more than others, however. Patients' interpersonal behavior has been studied the most, then motivation, and personality traits and coping styles and cognitive skills, all of which have been studied in more than ten separate psychotherapy datasets. These are assets considered universally important to all psychotherapy modalities, and are accordingly also studied across different therapies, except for cognitive skills which form the very foundation of cognitive and cognitive-behavioral therapies and have been mainly studied within these modalities. Focus, modulation of affects, flexibility of interaction with the interviewer, reflective ability, intelligence, and defense styles have been studied in 4-5 datasets. These are suitability criteria that have been emphasized since their early development within psychodynamic treatments, with the majority of the prediction studies also focusing on psychodynamic psychotherapies, with the exception of intelligence, mainly studied in cognitive therapy, in which its role is also highly stressed. Patients' self-concept and response to trial interpretation, on the

other hand, have both only been studied in one dataset, despite also having been considered highly relevant, if not the most important (Davanloo 1978), aspects of the evaluation of psychotherapy suitability. Overall suitability, taking account of several aspects of suitability, has so far been studied based on 5 datasets.

In long-term psychotherapy, however, only the prediction by patient's modulation of affects, interpersonal relationships, and overall suitability has been studied based on 1-2 datasets including psychodynamic and cognitive-behavioral therapies. The findings on these suitability aspects remain to be confirmed and all other suitability aspects are yet to be studied in long-term psychotherapies.

2.5.5.2 Issues related to the interpretation and comparability of suitability studies

The classification of psychological suitability measures is a challenging task as many of them are closely related. Poorer capacities regarding one aspect of suitability may also be balanced by superior capacities in another, and therefore evaluating one's overall suitability, instead of any one aspect alone, may be more relevant to the prediction of suitability for and the outcome of psychotherapy. This suggestion was also supported by the research evidence. Patients' overall suitability was found to be predictive of psychiatric symptoms in 4 out of 5 short-term psychotherapy datasets. Of the individual suitability measures, reflective ability, motivation, coping styles and cognitive skills, and interpersonal behavior were found to be associated with psychiatric symptoms in just over half of the short-term psychotherapy datasets in which they were studied (3/5, 8/14, 6/11, and 11/16, respectively). Modulation of affects, flexibility of interaction with the interviewer, and defense style were found to be predictive of psychiatric symptoms by half of the datasets (2/4). Intelligence and personality traits were associated with psychiatric symptoms in less than half of the datasets (2/5 and 3/11, respectively). The role of circumscribed focus as a predictor of psychiatric symptoms was not supported by any of the five datasets evaluating it, and the only study on self-concept and trial interpretation failed to find an association with psychiatric symptoms. Long-term psychotherapy datasets supported the role of modulation of affects and interpersonal relationships, but not overall suitability, as psychotherapy outcome predictors, but no replication studies were available.

Such a comparison of the suitability findings is, however, misleading in that several of the datasets included more than one suitability measure assessing the same aspect of suitability, several outcome measures, and several measurement points (see Table 6, Supplementary tables 1A-F). Taking into account all of the individual suitability results, only overall suitability was supported by the majority of results. Direct comparison of the existing prediction results is, however, unfair in that the sizes and characteristics of the patient samples, the types and lengths of the therapies and their follow-up, the statistical methods applied and adjustments for baseline symptoms and other confounding factors varied greatly. The majority of the

samples were small- or medium-sized and mainly included depressed women, undergoing a short-term, typically psychodynamic, cognitive or cognitive-behavioral therapy, with the outcome evaluation carried out only at the end of treatment and often without consideration of baseline symptoms or other confounding factors. Also, the suitability measures applied varied from one study to another, with very few replications of the same suitability measures. Established, well-validated, and readily available suitability measures (e.g., DAS, IIP), usually focusing on the evaluation of one dimension of suitability, were more widely applied and replicated; many of the suitability measures used, especially with respect to the nature of problems, ego strength, and self-observing capacity domains, were not, however, validated and/or available, thus prohibiting their replication. Both self-reported and interview-based suitability measures were used for the evaluation of most of the different aspects of suitability: only coping styles and cognitive skills were always evaluated via a self-report, and focus, flexibility of interaction with the interviewer, and the response to trial interpretation via interview, which for the latter two measures is a necessity. All of the measures of overall suitability applied evaluated at least one of these three aspects (Table 3), and were accordingly also all evaluated based on an interview. A variety of outcome measures were also used, but almost all of these were established, well-validated psychiatric symptom measures, commonly applied in psychotherapy research, with BDI, HDRS, and SCL-90-GSI being the most common measures. One fourth (11/41) of the patient samples were followed beyond the end of treatment; only three studies, however, followed the patients longer than one year (Husby et al. 1985, Husby 1985a; Høglend et al. 1992a, 1993, 1994, Høglend 1993; Solbakken et al. 2012) and none of these studies measured their symptom level at the end of the therapy. Little is thus known about the stability of the symptom level reached through therapy in the long term.

2.5.5.3 Prediction by suitability measures within short-term psychotherapy modalities

Of the different short-term therapy modalities studied, only interpretative PP and CBT were studied for most of the suitability measures: coping styles and cognitive skills was the only measure not studied in PP, and self-concept and the response to trial interpretation were the only measures not studied in CBT. Although most of these measures were first suggested as being relevant to suitability for interpretative PP, no research evidence supporting the role of the pre-treatment focus, modulation of affects, flexibility of interaction with the interviewer, self-concept, response to trial interpretation, motivation, defense style, or personality traits in the prediction of post-treatment symptoms in PP has yet been found. These suitability measures have, however, so far only been studied based on 1-3 datasets with relatively small sample sizes. The findings regarding reflective ability and interpersonal behavior have been contradictory, but manifesting insight and possessing good quality of relations have

been shown to be associated with reduced symptoms. The existing post-treatment follow-up findings supported the importance of modulation of affects, were contradictory regarding the role of flexibility of interaction, reflective ability, motivation, problem-solving capability, and interpersonal relationships, with some significant and some non-significant findings, and did not support prediction by focus, response to trial interpretation, defense style, and personality traits in the long run. Also, the findings regarding overall suitability were contradictory as only one of the two different studies applying the same suitability assessment scale, CDPS, found the total score to be significantly associated with post-treatment psychiatric symptoms, while another overall suitability score was not associated with post-treatment symptoms but was associated with symptoms one year after the end of treatment. The roles of reflective ability, defense style, and interpersonal behavior were also studied in supportive PP, generally considered useful for patients lacking capacities considered necessary for interpretative PP. The findings supported the positive role of a mature defense style, were contradictory regarding the role of the quality of interpersonal relations, with one significant and one non-significant finding, and did not support the relevance of reflective ability to symptom reduction in supportive PP. In CBT, no support for psychotherapy outcome prediction by focus and reflective ability was found, but prediction by modulation of affects, flexibility of interaction, defense style, and overall suitability was supported. All of these findings were, however, based on one single study (Myhr et al. 2007), with a sample size larger than in any of the studies on PP. Research evidence for CBT also supported the importance of some aspects of motivation (i.e., personal motivation, readiness and responsibility for change), coping styles and cognitive skills (i.e., lower cognitive dysfunction, lower perfectionism), personality (i.e., higher persistence, lower neuroticism), and interpersonal behavior (i.e., history of meaningful relationships) on positive psychotherapy outcome. CT was studied for five and BT for two suitability measures. Motivation, readiness to change, was associated with a better outcome in one CT study, lower intelligence with a worse outcome in one of the two intelligence studies, and interpersonal problems (i.e., general interpersonal distress, and underinvolved, overinvolved, avoidant, and ambivalent interpersonal style) with worse outcome in three out of four studies. The role of coping styles and cognitive skills and personality were not supported by the existing studies. In BT, both motivation and modulation of affects were found to be associated with significantly reduced post-treatment anxiety symptoms, but the latter finding did not persist in a further follow-up. IPT was studied for four suitability measures, and the findings supported the positive effect of good interpersonal behavior and personal motivation and the negative effect of some aspects of coping styles and cognitive skills (i.e., high perfectionism, dysfunctional attitudes) and personality traits (i.e., high harm avoidance, low self-directness, low novelty seeking, reward dependency, self-criticism) on psychotherapy outcome. ACT and ET were only studied for one suitability measure: high neuroticism was found to

predict a worse psychotherapy outcome in ACT, whereas no support for the role of interpersonal behavior in predicting the ET outcome was found.

2.5.5.4 Prediction by suitability measures between short-term psychotherapy modalities

Seven datasets included and compared several individual short-term psychotherapies. Interpretative PP was only compared with supportive PP in one study, which found patients with a better reflective ability and quality of interpersonal relations to benefit more from interpretative PP but not from supportive PP, supporting the ability of these criteria to separate those suitable for interpretative PP from those in need of supportive PP. CBT was compared with IPT in four studies with respect to the role of motivation, coping styles and cognitive skills, personality traits, and interpersonal behavior. No difference in prediction by motivation between these two therapies was found; one study found motivation to be predictive in both, and another study in neither. Regarding coping styles and cognitive skills, perfectionism and dysfunctional attitudes were found to predict a worse outcome both in CBT and IPT, whereas the need for approval was not found to be predictive in either. Lower cognitive dysfunction was, however, found to predict greater symptom reduction in CBT than in IPT, and lower social dysfunction greater symptom reduction in IPT than in CBT, supporting the cognitive foundation of CBT and interpersonal foundation of IPT. In line with this, another study found that patients with higher attachment avoidance showed greater symptom reduction in CBT than in IPT. Two studies comparing the psychotherapy outcome prediction by personality traits also showed that different personality traits may be harmful in CBT (i.e., low persistence) versus IPT (i.e., high harm avoidance, low self-directness, low novelty seeking, reward dependency, and self-criticism). One study, which did not find prediction by motivation or interpersonal relations to differ between CBT and IPT, combined the therapy groups in order to further study the prediction of these measures in a larger sample, but did not find support for prediction by either. CBT was also compared to BT in one study on intelligence and anger-hostility and to ACT in another study on neuroticism: the first study found no support for either measure in either therapy, whereas the latter study found higher neuroticism to be harmful in both.

2.5.5.5 Prediction by suitability measures within and between long-term psychotherapy modalities

Research evidence regarding the relevance of the psychological suitability measures to the prediction of the symptom outcome of long-term psychotherapies is only available for the modulation of affects, interpersonal relationships, and overall suitability in PP and CBT; all of these three aspects have been studied in PP but only interpersonal relationships in CBT. Patients with negative affective or interpersonal contributions to begin with, i.e., expressing impairment in their affect consciousness

and integration or hostility or hostile submission, controlling, or withdrawal in their interpersonal relations, were found to benefit more from PP. Also, positive interpersonal contributions were found beneficial to the PP outcome. Overall suitability, on the other hand, was not found to be predictive of the PP outcome. In the only study comparing PP and CBT, interpersonal problems were found predictive of the PP but not of the CBT outcome; patients with initially greater interpersonal problems were found to benefit more from PP. In general, long-term psychotherapies thus also seemed helpful for patients with poorer psychological capacities, as hypothesized. Replication of the existing findings and studies on the psychological suitability measures not yet studied in long-term psychotherapies are, however, needed.

2.5.5.6 Need for studies on the prediction by suitability measures in short-term versus long-term psychotherapy

The psychological suitability selection criteria were initially developed in order to facilitate the differentiation of those treatable with short-term psychotherapy from those in need of long-term psychotherapy. The literature review, however, showed that no studies comparing the suggested suitability criteria as predictors of psychotherapy outcome in short-term versus long-term psychotherapies have yet been published, and thus their hypothesized predictive validity is not yet known. Most of these criteria were first developed within short-term psychodynamic psychotherapies, but no studies comparing their prediction in short-term psychodynamic therapy versus other therapy modalities have been carried out, and prediction by half of the psychological suitability measures has not been compared between any therapy modalities; thus it is not known whether their prediction differs between different types of short-term therapies. Studies on suitability criteria as predictors of the outcome of individual short-term psychotherapies, including psychodynamic and other therapies, have been carried out but the design of the studies has varied considerably in terms of the choice of patients, therapies, suitability and outcome measures, as well as statistical methods. Studies on the suitability criteria as predictors of the outcome of individual long-term psychotherapies, including psychodynamic, are lacking for most of the suitability measures. The suitability studies carried out so far have mainly evaluated the symptom outcome rather than symptom profile, and studies on the long-term stability of the symptom outcome are scarce. Some of the important suitability selection criteria, i.e., self-concept and response to trial interpretation, have barely been studied and the few existing, validated suitability assessment scales (e.g., STAPP, CDPS, SSCT) do not include these criteria. There is thus an urgent need for standardization of the conduct of psychotherapy suitability studies and for filling the current gaps in the literature.

3 AIMS OF THE STUDY

In this study, a new interview-based 7-item psychotherapy suitability assessment scale, the Suitability for Psychotherapy Scale (SPS), addressing patients' psychological capacities that are suggested to predict and differentiate the outcome of short-term and long-term psychotherapy, was presented and studied for reliability, validity, and prediction. Prediction by the SPS of the outcome of short-term psychodynamic psychotherapy and solution-focused therapy was studied and compared for the first time, to determine whether the theoretical psychotherapy suitability criteria, first proposed within psychodynamic psychotherapies, work in practice and may also be applied to other psychotherapy modalities. Most importantly, prediction by the SPS of the outcome of these short-term therapies and long-term psychodynamic psychotherapy were compared for the first time, to determine whether these criteria may be applied when making a choice between short-term and long-term psychotherapy as originally suggested.

The more specific aims of the study, addressed in three substudies, are:

1. To present the SPS and assess its inter-rater reliability and repeatability, as well as its criterion and discriminating validity (Study I);

2. To determine and compare the prediction by the seven suitability measures of the SPS, alone and combined, of psychiatric symptoms in short-term psychodynamic psychotherapy and solution-focused therapy during a 3-year follow-up from the start of the therapies (Study II);

3. To determine and compare the prediction by the seven suitability measures of the SPS, alone and combined, of psychiatric symptoms in short-term, psychodynamic or solution-focused, therapy and long-term psychodynamic therapy during a 3-year follow-up from the start of the therapies (Study III).

These substudies are henceforth referred to by the above-mentioned Roman numerals.

4 POPULATION AND METHODS

4.1 Helsinki Psychotherapy Study (HPS)

This dissertation is based on data from the Helsinki Psychotherapy Study (HPS). HPS is a randomized clinical trial initiated to evaluate the effectiveness of solution-focused therapy (SFT), short-term psychodynamic psychotherapy (STPP), and long-term psychodynamic psychotherapy (LTPP) in the treatment of adult outpatients with mood or anxiety disorder (Knekt and Lindfors 2004, Knekt et al. 2008a). The study followed the Declaration of Helsinki and was approved by the ethics council of the Helsinki University Central Hospital. Written informed consent for participation was obtained from all patients.

4.1.1 Patients

Outpatients from the Helsinki area were referred to the study after psychiatric evaluation by local practitioners from June 1994 to June 2000 (Knekt and Lindfors 2004). Eligible patients were 20-45 years of age and had a long-standing (> 1 year) disorder causing dysfunction in work ability. The patients included also had to meet Diagnostic and Statistical Manual of Mental Disorders (DSM) criteria (American Psychiatric Association 1994; DSM-IV) for mood or anxiety disorders and criteria for neurosis to higher-level borderline personality organization (Kernberg 1996). Patients were excluded if they had a psychotic disorder or severe personality disorder (DSM-IV cluster A personality disorder and/or lower level borderline personality organization), bipolar I disorder, adjustment disorder, substance-related disorder, or severe organic disorder. Exclusion criteria further consisted of patients who had undergone psychotherapy within the previous two years, psychiatric health employees, and persons known to the research team.

A total of 459 patients were considered eligible (Figure 1), but 133 of them refused to participate. The study population thus consisted of 326 patients. Their mean age was 32 years and more than three quarters of them were female (Table 8, Studies I-III). More than half of the patients were living alone and one in four had completed a university degree. A total of 85% of the patients suffered from mood disorder, mainly major depressive disorder, and 44% from anxiety disorder, typically panic disorder or generalized anxiety disorder. Over one fourth of the patients thus had comorbid mood and anxiety disorder, and non-severe personality disorders were diagnosed in 18%. More than half of the patients had experienced their first psychiatric disorder at less than 22 years of age and major separations during childhood. About one in five patients had previously undergone psychotherapy or used psychotropic medication prior to baseline washout period.

Patient suitability for short-term
and long-term psychotherapy
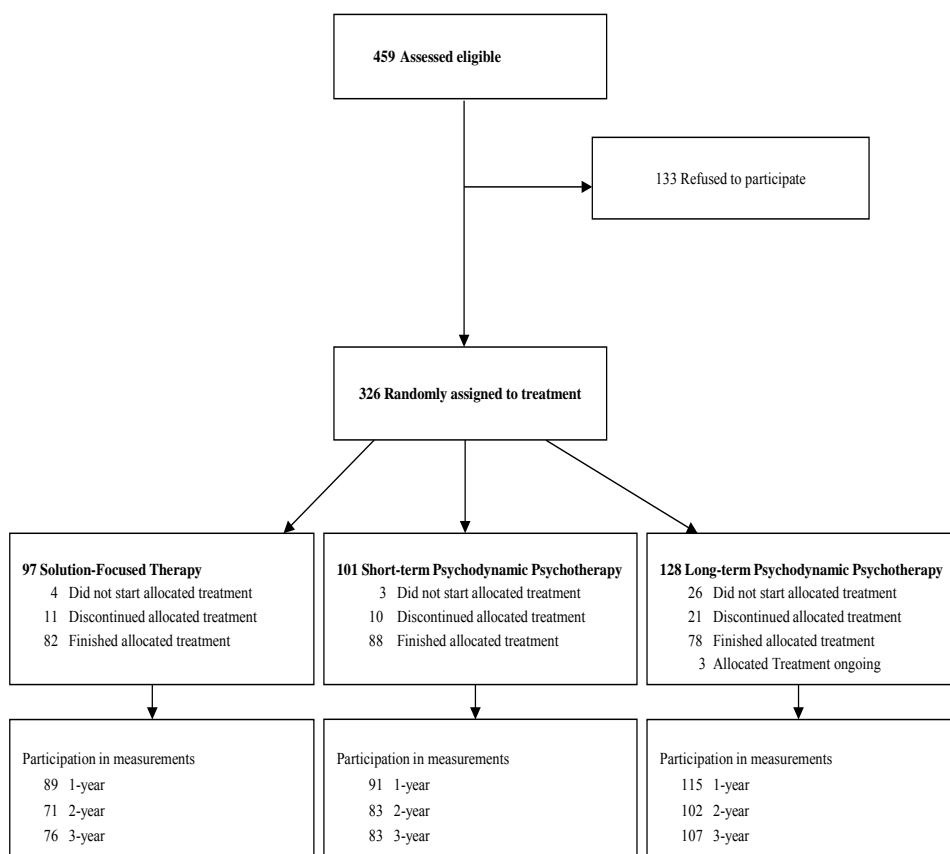
Figure 1. Number of eligible patients who were assigned to study group and completed the protocol.

The mean level of general psychiatric symptoms at baseline, according to Global Severity Index (GSI) of the Symptom Checklist-90 (SCL-90), was elevated with a mean value of 1.28 (Table 8). According to the Beck Depression Inventory (BDI) and Hamilton Depression Rating Scale (HDRS), the patients suffered from moderate depression with mean values of 18.3 and 15.7, respectively. The SCL-90 Anxiety scale (SCL-90-Anx) and Hamilton Anxiety Rating Scale (HARS) with values 1.24 and 14.9, respectively, suggested that the patients suffered from moderate anxiety symptoms.

These 326 patients were randomized, with the ratio 1:1:1.3, to SFT (N = 97), STPP (N = 101), or LTPP (N = 128). The randomization was successful in that no notable differences in the distribution of the patients' baseline characteristics across the three therapy groups were found (Table 8).

Table 8. Baseline characteristics of patients intended to treat with solution-focused therapy (SFT), short-term psychodynamic psychotherapy (STPP), or long-term psychodynamic psychotherapy (LTPP).

	Therapy group			
	SFT	**STPP**	**LTPP**	**ALL**
Variable	**(N = 97)**	**(N = 101)**	**(N = 128)**	**(N = 326)**
Sociodemographic variables				
Age (years) [1]	33.6 (7.2)	32.1 (7.0)	31.6 (6.6)	32.3 (6.9)
Men (%)	25.8	25.7	21.1	23.9
Living alone (%)	56.7	48.5	49.2	51.2
University degree (%)	28.9	19.8	28.1	25.8
Diagnoses				
Mood disorder (%)	86.6	78.2	88.3	84.6
Anxiety disorder (%)	46.4	49.5	36.7	43.6[2]
Comorbid mood and anxiety disorder (%)	33.0	27.7	25.0	28.2
Personality disorder (%)	18.6	24.8	12.5	18.1[2]
Psychiatric history				
Primary psychiatric disorder at age < 22 years (%)	66.0	57.6	63.0	62.2
Major separations during childhood (%)	44.3	51.5	50.8	50.9
Previous psychotherapy (%)	20.0	18.8	19.0	19.3
Previous psychotropic medication (%)	27.8	21.8	17.6	22.0
Symptoms				
Symptom Checklist-90, Global Severity Index (SCL-90-GSI)[1]	1.31 (0.50)	1.26 (0.53)	1.27 (0.55)	1.28 (0.52)
Beck Depression Inventory (BDI) [1]	18.1 (7.8)	17.9 (7.5)	18.8 (8.3)	18.3 (7.9)
Hamilton Depression Rating Scale (HDRS) [1]	15.8 (4.5)	15.4 (5.0)	15.8 (4.9)	15.7 (4.8)
Symptom Checklist-90, Anxiety scale (SCL-90-Anx)[1]	1.27 (0.72)	1.25 (0.66)	1.19 (0.68)	1.24 (0.69)
Hamilton Anxiety Rating Scale (HARS) [1]	14.9 (5.2)	15.0 (5.4)	14.8 (5.2)	14.9 (5.2)
Suitability for Psychotherapy Scale (SPS)				
Nature of problems				
Focus (%)[3]	39.2	34.0	36.7	36.6
Ego strength				
Modulation of affects (%)[3]	66.0	65.3	71.9	68.1
Flexibility of interaction (%)[3]	88.7	87.1	90.6	89.0
Self-concept in relation to ego ideal (%)[3]	81.4	80.2	85.2	82.5
Self-observing capacity				
Response to trial interpretation (%)[3]	74.2	64.4	64.8	67.5
Reflective ability (%)[3]	81.4	80.2	82.8	81.6
Motivation (%)[3]	39.2	38.6	39.1	39.0
SPS score (%)[3]	78.4	78.0	79.7	78.5

[1] $\bar{x}\,(SD)$

[2] Statistically significant (P < 0.05) difference between the short-term therapies and long-term therapy.

[3] Proportion of patients with good values in the suitability measure. See Table 9 and Appendix 3 for classification of the suitability measures.

4.1.2 Therapies

Solution-focused therapy (SFT) is a short-term resource-oriented and goal-focused therapeutic approach which helps clients change by constructing solutions (Johnson and Miller 1994, Lambert et al. 1998). The technique includes the search for pre-session change, miracle and scaling questions, exploration of exceptions, use of a one-way mirror and consulting break, positive feedback and home assignments. The orientation was based on an approach developed by de Shazer and Berg (de Shazer et al. 1986, de Shazer 1991). The frequency of sessions was flexible, usually once every second or third week, up to a maximum of 12 sessions, over no more than 8 months.

Short-term psychodynamic psychotherapy (STPP) is a short-term, focal, transference-based therapeutic approach which helps patients by exploring and working through specific intrapsychtic and interpersonal conflicts. STPP is characterized by the exploration of a focus, which can be identified by both the therapist and the patient. This consists of material from current and past interpersonal and intrapsychic conflicts and the application of confrontation, clarification, and interpretation in a process in which the therapist is active in creating the alliance and ensuring the time-limited focus. The orientation was based on approaches described by Malan (1976a) and Sifneos (1978). The therapy was scheduled for 20 treatment sessions, one session a week, over 5–6 months.

Long-term psychodynamic psychotherapy (LTPP) is an open-ended, intensive, transference-based therapeutic approach which helps patients by exploring and working through a broad area of intrapsychic and interpersonal conflicts. LTPP is characterized by a framework in which the central elements are exploration of unconscious conflicts, developmental deficits, and distortions of intrapsychic structures. Confrontation, clarification, and interpretation are major elements, as well as the therapist's actions in ensuring the alliance and working through the therapeutic relationship to attain conflict resolution and greater self-awareness. Therapy includes both expressive and supportive elements, the use of which depends on patient needs. The orientation followed the clinical principles of long-term psychodynamic psychotherapy (Gabbard 2004). The frequency of sessions was 2-3 sessions a week and the duration of therapy up to 3 years.

SFT was manualized and centralized clinical adherence monitoring of supervised cases was performed. Both psychodynamic psychotherapies were conducted in accordance with clinical practice, where the therapists might modify their interventions according to the patients' needs within the respective framework. Accordingly, no manuals were used and no adherence monitoring was organized.

After randomization, participation was refused by 33 patients (4 assigned to SFT, 3 to STPP, and 26 to LTPP), and 42 of the 293 patients starting the treatment discontinued prematurely (11 in SFT, 10 in STPP, and 21 in LTPP group) (Figure 1). The main reasons for refusal were objection to the type of psychotherapy

assigned to the patient or patient's life situation, and the main reasons for discontinuation were disappointment with the treatment or life situation (Knekt et al. 2008a). Patients discontinuing SFT had more symptoms than those continuing treatment. The average number of therapy sessions among patients starting therapy was 10 (SD = 3.3) for SFT, 19 (SD = 3.4) for STPP, and 232 (SD = 105) for LTPP. The observed mean lengths of therapies were 7.5 months (SD = 3.0) for SFT, 5.7 months (SD = 1.3) for STPP, and 31.3 months (SD = 11.9) for LTPP.

4.1.3 Therapists

The therapies were carried out by 55 therapists, of whom 6 provided SFT, 12 STPP, and 41 LTPP (Knekt et al. 2008a, Heinonen et al. 2012). Eligible therapists were required to have at least two years' experience in the respective therapy form after completion of their training. The therapists providing STPP and LTPP were mainly psychologists (83% and 81%, respectively) whereas those providing SFT had a more heterogeneous educational background (e.g., psychologists, physicians, or social workers). The therapists who provided SFT had been trained in the method and had received a qualification in SFT from a local accredited institute. The therapists who provided STPP or LTPP had received standard training in psychoanalytically oriented psychotherapy by one of the accredited psychoanalytic or psychodynamic training institutes in Finland. Training adhered to clinical principles of psychodynamic orientation and technique although the emphasis of different theoretical models varied (e.g., ego psychological, object-relations, self-psychological, and attachment models) (Gabbard 2004). All psychodynamically oriented therapists had completed a minimum of 3 years' analytical training in psychodynamic psychotherapy (psychoanalysis or long-term psychodynamic psychotherapy), and those providing short-term therapy had completed an additional 1-2 years of specific short-term focal psychodynamic therapy training. None of the psychodynamic therapists had any experience of SFT and vice versa. The mean number of years of experience in the therapy form provided after completion of clinical training was 9 (range = 3-15) in SFT, 9 (range = 2-20) in STPP, and 18 (range = 6-30) in LTPP. The therapists providing STPP had, in addition, on average 16 (range = 10-21) years of experience in LTPP.

4.2 Study design and assessment methods

The present study was carried out as a cohort study with repeated measurements: the patients were assessed at baseline and after 3, 7, 9, 12, 18, 24, and 36 months of follow-up. During this 3-year follow-up after randomization to the therapies, the patients were provided with either short-term therapy, followed by no treatment, or long-term therapy (Knekt et al. 2008a).

4.2.1 Suitability for Psychotherapy Scale (SPS)

4.2.1.1 Development and content of the SPS

Patients' psychological capacities, considered to reflect patient suitability for psychotherapy (see Section 2.2), were assessed at baseline using a new interview-based 7-item Suitability for Psychotherapy Scale (SPS), developed by a clinician and psychotherapy trainer Veikko Aalberg and his colleagues (Appendix 3, Studies I-III). Its development was influenced by the international pioneers of psychotherapy suitability Michal Balint and David Malan (Balint 1972, Malan 1976a), and Peter Sifneos (Sifneos 1979), and the Finnish pioneer of psychoanalysis, psychodynamic psychiatry, and psychotherapy Veikko Tähkä. Aalberg and his colleagues organized short-term psychotherapy training and carried out patient psychotherapy suitability evaluations over several years, much like Balint, Malan, and Sifneos, to test and refine their suitability criteria which ultimately led to the development of the SPS. The SPS covered demarcation of focus, i.e., central theme or conflict underlying patient's current distress, and assessment of patient's modulation of affects, flexibility of interaction with the interviewer, self-concept in relation to ego ideal, response to trial interpretation, reflective ability, and motivation. These seven suitability measures were considered to predict and differentiate suitability for and accordingly the outcome of short-term versus long-term psychotherapy, with good values in these measures serving as indications for short-term psychotherapy and poor values as contraindications for short-term, but not for long-term, psychotherapy.

The seven individual measures of the SPS can be further classified under three suitability domains – nature of problems, ego strength, and self-observing capacity – according to their clinically relevant conceptual scope (see Table 1).

Nature of problems

The nature of problems domain comprised the measure of the focality of the problems (Malan 1976a, Sifneos 1979).

Determination of circumscribed focus referred to the individual's ability to identify mentally represented basic conflicts related to his or her key problems and the ability to consider them as dynamic, etiological components of his or her current difficulties. The focus was considered to be circumscribed when one's problem area was experientially presented, clearly defined, and restricted in scope. The focus was not considered to be circumscribed when it was not presented experientially and when the problem area was global or unspecific, or remained undefined.

Ego strength

The ego strength domain included measures representing psychological capacities related to dealing with affects, interpersonal relations, and self-structure, all being derivatives of the global psychodynamic concept of ego strength and expressing thus

different aspects of the capacity to deal adaptively with internal and external demands and reality (Kernberg et al. 1972, Lake 1985).

Modulation of affects was defined as the individual's ability to regulate and express both positive and negative feelings. Adequate modulation of affects was indicated by a realistic recognition of one's emotions and by a good ability for affect control. Restricted contact with affects or uncontrolled affects, on the other hand, was considered to indicate inadequate modulation of affects.

Flexibility of interaction was defined as the individual's ability to establish a good working dialog and collaboration in working towards the goals set out by the interviewer with the interviewer. Good or fairly good dialog without major restrictions or difficulties in the collaboration indicated good or fair interaction. Correspondingly, when the dialog was significantly restricted and dysfunctional, the flexibility of interaction was considered to be poor.

Self-concept in relation to ego ideal was defined as a form of narcissistic balance as manifested in the relation between the individual's current self-concept, expectations directed to oneself, and his or her abilities. The relationship between self-concept and ego ideal was considered balanced when the self-concept was realistic and when only minor belittlement of self appeared and when ego ideal was attainable. Moderately disturbed relationships included more severe self-denigration and unrealistic features in self-concept or ego ideal. A severely disturbed relationship was evident when both self-denigration and unrealistic or grandiose features in an ideal self were observed.

Self-observing capacity

The self-observing capacity domain consisted of measures related more to an orientation towards therapy and capacities for the process, including response to trial interpretation, reflective ability, and motivation for addressing problems psychologically, extending and integrating elements of psychological mindedness and motivation for change (Appelbaum 1972, Piper et al. 1998).

The nature of trial interpretation and the individual's response to it was defined based on the recognition of the problem area and the individual's ability to elaborate on the theme of the interpretation made by the interviewer. A good response to the trial interpretation was characterized by an experience-near elaboration of the interpretation. A fair response was characterized by a delayed or missing elaboration with non-experiential material. A poor response was evident when there was a total lack of additional material offered to support or negate the interpretation, or when the interviewer was unable to make any interpretations due to lack of material.

Reflective ability was defined as an emotional-cognitive ability to perceive and recognize one's hopes and desires and the ability to formulate links between the present and earlier phases of life in a psychologically meaningful way. Very good reflective ability was evident when there was elaborated psychological introspection and consideration of the time span available in the interview. Good or fair reflective

ability was indicated when there was less elaboration or more defensiveness in introspection but a consideration of the time span. Reflective ability was considered restricted when it was limited to only the present situation or external facts and when there was no psychological elaboration at all.

Motivation was defined as the individual's willingness and commitment to address problems psychologically. Good motivation was indicated by the individual's subjective desire to seek therapy and to work introspectively. Ambivalent or restricted motivation was based more on seeking therapy on someone else's recommendation with non-introspective aims, or being otherwise reserved or ambivalent towards the therapy.

4.2.1.2 Evaluation of the SPS

The assessment of suitability for psychotherapy was based on three initial interview sessions, with a minimum of 90 minutes needed for the specific assessment of the suitability measures, conducted by trained and experienced clinical interviewers, who were not involved in the patients' treatment (Study I, Knekt and Lindfors 2004). The first two interview sessions (both 45 minutes) followed a semi-structured procedure based on a modification of Kernberg's Structural Interview (Kernberg 1981). They covered the standard elements of a psychiatric evaluation, including the case history and psychiatric diagnosis, as well as the assessment of the suitability measures (Appendix 3). Most of the third interview session (60 minutes) was structured and focused on diagnostic evaluation and the assessment of several structured symptom measures.

During the semi-structured phase, the interviewer presented several opening questions for the patient regarding the predominant complaints that had made him or her seek therapy. After that, the interviewer focused on observing the patient's capacity to elaborate on the theme, offering clarifications and making confrontations if necessary and evaluating the patient's ability to utilize them. If the patient produced no information or only very limited information in a certain key area, the interviewer commented on it (e.g., 'You spoke very little about your mother') to elicit a more detailed description indirectly. The aim was to obtain a comprehensive view regarding the psychological functioning and capacities and personal history of the patient. Evaluation of the here-and-now affective and reflective functioning and identity issues, as well as the interaction with the interviewer, was essential in the procedure.

4.2.1.3 Classification of the SPS measures and formation of the SPS score

Each of the seven suitability measures was assessed on a 7-point scale from 1 to 7, where low and intermediate values indicated good suitability and high values poor suitability, except for focus and motivation, for which both intermediate and high values were considered to indicate poor suitability (Table 9, Study I). An SPS score was formed by summing up the values of the seven dichotomized suitability

Table 9. Description and classification of the seven individual suitability measures of the Suitability for Psychotherapy Scale (SPS) and the cumulative SPS score.

Suitability measure	N (%)[1]	3 classes[2]	2 classes[3]
DOMAIN 1: Nature of problems			
Focus			
1: Circumscribed and experiential focal conflict present at first interview	39 (12.0)		
2: Circumscribed and experiential focal conflict present at 2nd or 3rd interview	80 (24.6)	Low	Good (0)[4]
3: Specific focal conflict with no experiential derivative	46 (14.2)		
4: Global problem area and theoretically derived focal conflict with no clinical vignette	115 (35.4)	Intermediate	
5: Unspecific and mainly theoretical focus, e.g., dependency problem	42 (12.9)		
6: Very difficult to determine a focus for short-term therapy	3 (0.9)		
7: Impossible to determine any focus	0 (0.0)	High	Poor (1)[4]
DOMAIN 2: Ego strength			
Modulation of affects			
1: Contact with both positive and negative affects, good ability for adequate affect control	26 (8.0)		
2: Mild defensiveness, mainly good modulation of affects	79 (24.2)	Low	
3: Somewhat defensiveness without major impact on the interview, restricted contact with affects	117 (35.9)	Intermediate	Good (0)[5]
4: Significant defensiveness, narrowing the contact with interviewer	62 (19.0)		
5: Very pronounced defensiveness or moderate affective outbursts	38 (11.7)		
6: Disaffected or affective outbursts	4 (1.2)		
7: Affective stupor or agitation	0 (0.0)	High	Poor (1)[5]
Flexibility of interaction			
1: Very good, flexible and natural dialogue	21 (6.4)		
2: Good, mild difficulty in collaboration or exceeding time limits	167 (51.2)	Low	
3: Fair, temporary breaks of narration	70 (21.5)		
4: Restricted, does not take notice of time limits	32 (9.8)	Intermediate	Good (0)[5]
5: Significantly restricted, nonintegrated separate themes	27 (8.3)		
6: Dysfunctional dialogue and loss of time perspective	8 (2.5)		
7: Very poor, whole interview dysfunctional	1 (0.3)	High	Poor (1)[5]
Self-concept in relation to ego ideal			
1: Adequate balance between self and ideal self	8 (2.5)		
2: Belittling of self, ideal self attainable	46 (14.1)	Low	
3: Self denigration, ideal self difficult to attain	135 (41.4)		
4: Some unrealistic features in self and ideal self	80 (24.6)	Intermediate	Good (0)[5]
5: Self denigration and grandiose features in ideal self	50 (15.3)		
6: Unrealistic and grandiose features in self structure and ideal self	4 (1.2)		
7: Denigration of self and grandiose ideal self demands, leading to stagnation	3 (0.9)	High	Poor (1)[5]

Suitability measure	N (%)[1]		3 classes[2]	2 classes[3]
DOMAIN 3: Self-observing capacity				
Response to trial interpretation				
1: Experience-near elaboration at first interview	9	(2.8)		
2: Experience-near elaboration at 2nd or 3rd interview	31	(9.5)	Low	
3: Non-experiential or delayed elaboration	53	(16.2)		
4: Responsiveness with extra material and no elaboration	127	(39.0)	Intermediate	Good (0)[5]
5: Responsiveness with no extra material and no elaboration	80	(24.5)		
6: No responsiveness to trial interpretation	16	(4.9)		
7: No trial interpretation made by the interviewer	10	(3.1)	High	Poor (1)[5]
Reflective ability				
1: Very good, psychological elaboration and consideration of time span	15	(4.6)		
2: Good, mild restrictions, consideration of time span	134	(41.1)	Low	
3: Fair, defensive narrowness of reflective ability, consideration of time span	117	(35.9)	Intermediate	Good (0)[5]
4: Restricted to external facts, consideration of time span	29	(8.9)		
5: Significantly restricted, narrow time span	22	(6.7)		
6: Severe restriction to only present situation	9	(2.8)		
7: No psychological reflective ability	0	(0.0)	High	Poor (1)[5]
Motivation				
1: Very good, based on own activity and long-term consideration, desire for introspection	24	(7.4)		
2: Good, based on own activity, introspective needs less prominent	103	(31.6)	Low	Good (0)[4]
3: Good, problem relief main reason for seeking treatment	142	(43.6)		
4: Quite good, problem relief almost the only reason for seeking treatment	32	(9.8)	Intermediate	
5: Fair, ambivalent, based significantly on others' suggestions	21	(6.4)		
6: Poor, significant ambivalente	4	(1.2)		
7: No motivation for therapy	0	(0.0)	High	Poor (1)[4]
SPS score[6]				
0: All 7 suitability measures good	40	(12.3)		
1: 6 good, 1 poor	72	(22.1)		
2: 5 good, 2 poor	88	(27.0)		
3: 4 good, 3 poor	56	(17.2)	Low	Good (0)[4]
4: 3 good, 4 poor	27	(8.3)		
5: 2 good, 5 poor	22	(6.8)		
6: 1 good, 6 poor	13	(4.0)	Intermediate	
7: All 7 suitability measures poor	8	(2.5)	High	Poor (1)[4]

[1] Division of 326 patients to different values (1-7) of the seven original suitability measures of the SPS and to different values (0-7) of the cumulative SPS score.

[2] Classification of the values (1-7) of the 7 individual suitability measures of the SPS and the cumulative SPS score (0-7) into low, intermediate, and high values.

[3] Classification of the low, intermediate, and high values of the suitability measures into good and poor.

[4] Only low values were considered good, whereas both intermediate and high values were considered poor.

[5] Both good and intermediate values were considered good, whereas high values were considered poor.

[6] Sum of good (0) or poor (1) values of the seven individual suitability measures varies between 0-7.

measures (good suitability = 0, poor suitability = 1) so that the score varied from 0 to 7. The SPS score was further analyzed as low (0-3), intermediate (4-6), or high (7), where low values were considered to indicate good overall suitability (patient had more good than poor values) and intermediate and high values poor overall suitability (patient had more poor than good values).

The majority (68%-89%) of the 326 patients in this study had good values in the individual SPS measures (Tables 8 and 9). The only exceptions were focus and motivation for which only low values (39% and 37%, respectively) were considered to indicate good suitability. According to the cumulative SPS score, almost 80% of the patients had good overall suitability. No notable differences in the distribution of patients with good and poor values in the suitability measures across the three therapy groups were observed (Table 8).

4.2.1.4 Validation of the SPS

To control the quality of the suitability assessment interviews and evaluate the reliability of the SPS assessments, a non-random sample of 37 videotaped initial and follow-up assessment interviews of 26 patients was selected (Study I). The sample was further completed with a videotaped interview of one patient with no symptoms of psychiatric disorder and one psychotic patient, in order to make it represent a socio-economically, clinically, and diagnostically heterogeneous group of patients. A total sample thus comprised 39 videotaped interviews of 28 patients. These 28 patients represented well the study sample of 326 patients (Study I).

Seven raters assessed the seven suitability measures based on the 39 videotaped interviews. Each rater repeated their suitability assessments after 3 years for 27 of the 39 videotaped interviews. These assessments were then used to determine both the agreement between the individual raters' suitability assessments and the repeatability of the individual raters' assessments.

The ratings were carried out by seven clinical interviewers (five psychologists and two psychiatrists), with a good deal of experience (range 9-20 years) in clinical practice and psychotherapy evaluation. In addition, separate training for the evaluation of the SPS was carried out in two phases over a period of several months. The first phase included four raters and was carried out between 1994 and 1997, with an intensive training phase of approximately 100 hours. The second phase for the last three raters, who joined the project later, was carried out during the first six months in 1998 and included an intensive 4-month group training of approximately 60 hours.

4.2.2 Psychotherapy outcome

The psychotherapy outcome was assessed via psychiatric symptoms, self-reported by the patients (Studies II and III). The psychiatric symptoms were assessed at baseline and 7 times (3, 7, 9, 12, 18, 24, and 36 months after baseline) during the 3-year follow-up, using the validated Finnish translation (Holi et al. 1998) of the

validated English Symptom Checklist-90 (SCL-90) questionnaire (Derogatis et al. 1973, 1976, Derogatis 1983, 2000). The SCL-90 consists of 90 items describing different problems and complaints, which were scored on a 5-point Likert scale of distress from 0 (none) to 4 (extreme), indicating how much discomfort that problem had caused the patient during the past month. The SCL-90 usually requires 12-20 minutes to complete (Derogatis 2000). Subsets of the 90 items (6-13 items) form basis for 9 symptom scales yielding scores on 9 primary symptom dimensions: somatization, obsessive-compulsive disorder, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, psychotism. The Global Severity Index (GSI), which is the average score of the 90 items of the questionnaire, is the summary of the test, designed to measure overall psychological distress. As the study population in this study comprised both patients suffering from mood and anxiety disorder, the SCL-90-GSI was chosen as a measure of their global symptom severity.

A total of 20 patients (7 in SFT, 2 in STPP, and 11 in LTPP) did not complete the outcome assessment at any other measurement occasion except for baseline measurement. The participation to outcome assessment varied between 67%-98% in SFT, 76%-99% in STPP, and 72%-98% in LTPP from one measurement point to another. The mean participation rates in SFT, STPP, and LTPP during the 3-year follow-up were 81%, 84%, and 79%, respectively.

4.2.3 Potential confounding factors

The factors that potentially confound the relationship between the suitability measures and psychotherapy outcome, and thus need to be adjusted for in the statistical analysis (Rothman et al. 2008), were assessed at baseline based on both interviews and self-report questionnaires (Studies II and III). Psychiatric diagnoses at Axes I and II were assessed based on a semi-structured diagnostic interview (Knekt and Lindfors 2004) according to the DSM-IV diagnostic criteria (American Psychiatric Association 1994), and criteria for neurosis to higher-level borderline personality organization based on a psychodynamic assessment interview (Kernberg 1996). Information on socio-demographic factors (sex, age, marital status, and education) and psychiatric history (age at the onset of first psychiatric disorder, separation experiences at childhood, and previous psychiatric treatment (i.e., psychotherapy, psychotropic medication, and hospitalization)) were assessed via interviews and questionnaires. Depressive and anxiety symptoms were both assessed based on interviews, using the 17-item Hamilton Depression Rating Scale (HDRS; Hamilton 1960) and the 14-item Hamilton Anxiety Rating Scale (HARS; Hamilton 1959), and self-report questionnaires, using the 21-item Beck Depression Inventory (BDI; Beck et al. 1961) and the 10-item Symptom Checklist-90, Anxiety Scale (SCL-90-Anx; Derogatis et al. 1973). Finally, a variety of personality functions were assessed using both interviews (the Level of Personality Organization (LPO;

Kernberg 1996), the Quality of Object Relations Scale (QORS; Azim et al. 1991)) and questionnaires (the 88-item Defense Style Questionnaire (DSQ; Andrews et al. 1989), the 36-item Structural Analysis of Social Behavior (SASB) introject questionnaire (Benjamin 1996), the 127-item Inventory of Interpersonal Problems (IIP; Horowitz et al. 1988), and the 29-item Sense of Coherence Scale (SOC; Antonovsky 1993)).

4.2.4 Auxiliary treatment

As an indicator of sufficiency of the treatments given, information on the use of auxiliary treatment (i.e., psychotherapy other than assigned in the treatment protocol, psychotropic medication, and psychiatric hospitalization) during the 3-year follow-up was continuously assessed by questionnaires, interviews, and from nationwide public health registers (Knekt and Lindfors 2004, Knekt et al. 2011b). Incomplete information was complemented by telephone inquiries.

In accordance with the inclusion criteria, none of the patients were in therapy or hospitalized at baseline, whereas a total of 20% of the patients used psychotropic medication before washout preceding baseline evaluation (Knekt et al. 2011b). A total of 40% of the patients used some type of auxiliary treatment, mainly psychotropic medication or auxiliary psychotherapy, during the 3-year follow-up. Auxiliary treatment, especially auxiliary psychotherapy, was more common in the short-term therapy groups (45% in SFT and 48% in STPP) than in the long-term therapy group (31% in LTPP).

4.3 Statistical methods

4.3.1 Validation of the SPS (Study I)

The reliability of the SPS assessments was evaluated by measuring both the agreement between the individual raters' assessments (N = 39) and the repeatability of the individual raters' assessments over 3 years (N = 27) by using the weighted kappa coefficient (Fleiss 1981). Agreement was evaluated poor, fair, good, and excellent for kappa values 0.00-0.39, 0.40-0.59, 0.60-0.74, 0.75-1.00, respectively. The significance of differences between the assessments was tested with the symmetry test.

The validity of the SPS assessments was evaluated at baseline in the total sample (N = 326) by measuring both criterion validity and discriminating validity. The criterion validity was evaluated by measuring whether there was a significant association between the SPS measures and a criterion measure, the Quality of Object Relations Scale (QORS; Azim et al. 1991), measuring theoretically similar phenomenon (suitability). The discriminating validity, on the other hand, was evaluated by measuring whether there was a lack of association between the SPS measures and a discriminating measure, the Symptom Checklist-90, Global Severity

Index (SCL-90-GSI; Derogatis et al. 1973), measuring theoretically different phenomenon (symptoms). A linear regression model was used in the estimation of the association between the SPS measures and the criterion and discriminating measures (Searle 1971, Cohen and Cohen 1975). A test for trend was performed by including the independent variable (SPS measures) as a continuous variable in the model.

4.3.2 Psychotherapy outcome prediction by the SPS (Studies II and III)

The main analyses on prediction by the SPS measures on psychotherapy outcome were based on the 'intention-to-treat' (ITT) design, in which patients are analyzed according to their initial treatment assignment in order to evaluate the clinical effect of the treatment policy. Complementary 'as-treated' (AT) analyses, which consider the compliance with the treatment assignment and aim at approximating the true treatment efficacy, were also performed (Härkänen et al. 2005, Knekt et al. 2008a). The primary analyses were based on the assumption of ignorable dropouts (Knekt et al. 2008a). In the secondary analyses, missing values were replaced by multiple imputation using Markov chain Monte Carlo methods (Rubin 1987). The statistical analyses were based on linear mixed models (Verbeke and Molenberghs 1997). Model-adjusted outcome means and mean differences were calculated for different measurement points (Lee 1981). The delta method was applied to calculate confidence intervals (Migon and Gamerman 1999). Statistical significance was tested with the Wald test.

In the ITT analyses, two models were used: a basic model and a complete model. The dependent variable in all analyses was the symptom measure SCL-90-GSI. The basic ITT model included as independent variables the SPS measure, therapy group, and time (i.e., follow-up measurement points), their first- and second-order interactions, and a correction term (i.e., the first-order interaction of the difference between theoretical and realized date of measurement, time, and the SPS measure). The independent variable of main interest was the interaction term between the SPS measure, therapy group, and time, telling us whether the association between the baseline level of the suitability measure and symptoms at follow-up differed between therapy groups compared and how. Separate models for each of the seven individual dichotomized SPS measures and the cumulative SPS score, representing overall suitability, were carried out. The two short-term therapy groups (SFT and STPP) were first analyzed separately to reveal any potential differences in the prediction by the SPS measures between them (Study II), but as no notable differences were found, they were combined into one short-term therapy group which was compared to the long-term therapy group (LTPP) (Study III). Since the patients were not randomized with respect to the SPS, the complete ITT model further included all the variables that satisfied the criteria for confounding, i.e., were associated with the SPS measures and the symptom outcome, without being an

intermediate or latent variable (Rothman et al. 2008): age, sex, marital status, education, DSM-IV diagnoses (comorbidity of mood and anxiety disorder, personality disorder), age at the onset of the first psychiatric disorder, separation experiences at childhood, previous use of psychotherapy or psychotropic medication, psychiatric symptoms (HDRS, HARS, BDI, SCL-90-Anx), and personality functions (LPO, QORS, DSQ, SASB, IIP, SOC). Both an unadjusted model and a model adjusted for the outcome variable at baseline were conducted.

To further account for the deviations from the study protocol, an AT model adjusting for non-compliance was created by adding variables describing compliance as baseline variables (waiting time from randomization to initiation of treatment and withdrawal before start of treatment) and as time-dependent covariates (discontinuation of study treatment and auxiliary treatment, i.e., psychotherapy, psychotropic medication, and psychiatric hospitalization, during follow-up) as main effects to the complete ITT model.

All three models (ITT basic, ITT complete, AT) were carried out based on both the original data and imputed data. Although imputation attenuated the results obtained based on the original data, the statistical significance of the findings remained and thus the conclusions were not changed (data not shown). No noticeable differences between the three different models were found. For these reasons, the results based on the original data and the simplest basic ITT model will be presented.

The significance of the SPS measures in predicting the outcome of the two short-term therapies (Study II) or short-term versus long-term therapy (Study III) during the 3-year follow-up period was determined based on the following three criteria. First, the statistical significance of the interaction between the SPS measure and the therapy group throughout the follow-up was tested using the Wald test. Second, the difference in the outcome between the therapy groups within categories of the SPS measure (good and poor) at the different measurement points was determined and tested for difference from zero. Third, the statistical significance of the change in outcome from the baseline to the different measurement points for each therapy group and category of the SPS measure was assessed.

The following inferences were drawn regarding the suitability of the two short-term therapies (Study II) or short-term and long-term therapy (Study III), based on the second and third criterion. Both short-term therapies or short-term and long-term therapy were considered equally beneficial when no statistically significant differences in patients' outcome between the therapy groups were found and when patients in both groups experienced and maintained a statistically significant reduction in symptoms in comparison to the baseline symptom level during the 3-year follow-up period. If patients benefited statistically significantly more from one of the therapies, or if patients experienced and maintained a statistically significant reduction in symptoms in comparison to the baseline symptom level only in one of the therapy groups, that therapy was considered more beneficial than the other.

Finally, if patients did not experience an enduring statistically significant symptom reduction in either of the therapies, neither of them was considered sufficient.

To allow comparison of the results from this study with the results previously presented in the literature, mainly based on correlation between the suitability measures and psychotherapy outcome, as well as to allow comparison between less and more advanced statistical methods, correlation coefficients for the association between the SPS measures and SCL-90-GSI were also provided.

The statistical analyses were carried out using SAS software, version 9.2. (SAS Institute Inc. 2008).

5 RESULTS

5.1 Validation of the SPS (Study I)

The quality of the suitability assessments based on the new interview-based 7-item Suitability for Psychotherapy Scale (SPS) was evaluated by measuring their reliability and validity (Study I). Reliability was evaluated by measuring both the agreement between individual raters' suitability assessments and the repeatability of the individual raters' assessments over time. Validity was evaluated by measuring both criterion and discriminating validity.

5.1.1 Reliability of the SPS

The agreement between the seven individual raters' suitability assessments was fair to good for most of the suitability measures, with a median kappa coefficient over interviewers for the 7-category suitability measures varying from 0.42 to 0.65 (Table 10). The only exceptions were agreement on patient's response to trial interpretation and motivation, with respective median kappa coefficients of 0.28 and 0.29. The kappa coefficients for agreement between the individual raters varied from 0.00 to 0.84, with the largest variation in trial interpretation (0.00-0.62) and motivation (0.00-0.58). There were no statistically significant systematic differences between the assessments made by different raters. The median kappa coefficient for agreement regarding the SPS score was 0.54.

The kappa coefficients representing the repeatability of the individual raters' assessments made three years apart showed nearly fair to good agreement beyond chance (median kappa over interviewers 0.36-0.59; Table 10). The most considerable variation in repeatability of the seven individual raters' assessments was observed for focus and response to trial interpretation, with kappa coefficients varying from 0.00 to 0.67 and 0.19 to 0.79, respectively. No systematic differences between the two measurement occasions were found, with the exception of one interviewer's evaluation of one patient's motivation. The median kappa coefficient for repeatability of the SPS score was 0.60.

5.1.2 Validity of the SPS

A significant relationship was hypothesized between the suitability measures and the criterion measure QORS, and no association between the suitability measures and the discriminating measure SCL-90-GSI. The correlation coefficients between the suitability measures and QORS varied from 0.18 (focus) to 0.46 (self-concept in relation to ego ideal) (Table 11), and a statistically significant inverse association was found (as smaller suitability measure values and higher QORS values are better). The correlation between the SPS score and QORS was 0.32 (p-value for

Patient suitability for short-term
and long-term psychotherapy

Table 10. Reliability of the Suitability for Psychotherapy Scale (SPS): Weighted kappa coefficients for agreement between seven individual raters' SPS assessments (N = 39) and for repeatability of the individual raters' SPS assessments over 3 years (N = 27).

Suitability measure	Agreement between raters			Repeatability of raters		
	N	Median[1]	Min-Max	N	Median[1]	Min-Max
Nature of problems						
Focus	19-20	0.42	0.08-0.62	10	0.38	0.00-0.67
Ego strength						
Modulation of affects	37-39	0.63	0.46-0.78	27	0.53	0.39-0.67
Flexibility of interaction	37-39	0.65	0.39-0.84	26-27	0.59	0.39-0.69
Self-concept	37-39	0.45	0.20-0.64	27	0.36	0.23-0.55
Self-observing capacity						
Trial interpretation	19-20	0.28	0.00-0.62	10	0.51	0.19-0.79
Reflective ability	37-39	0.59	0.42-0.72	27	0.45	0.36-0.62
Motivation	34-38	0.29	0.00-0.58	24-26	0.43	0.30-0.56[2]
SPS score	19-20	0.54	0.27-0.74	9-10	0.60	0.19-0.80

[1] Median over weighted kappa coefficients of the seven raters' assessments.

[2] The evaluations of the same patient made by the same rater 3 years apart differed from one another statistically significantly for one rater's one rating. Tested using the test of symmetry.

trend < 0.001). By contrast, the association of the suitability measures with SCL-90-GSI was much weaker, with correlation coefficients varying from 0.06 to 0.22, and a significant trend only found for two suitability measures (focus and self-concept in relation to ego ideal). A weak association was observed at baseline between the SPS score and SCL-90-GSI (r = 0.16, p-value for trend = 0.93).

5.2 Associations among the suitability measures and between the suitability measures and the outcome

The correlation matrix revealed both strong intra- and intercorrelations for different suitability domains (Table 12). Strong correlations were found within the ego strength domain between modulation of affects and flexibility of interaction (r = 0.48), and within the self-observing capacity domain between reflective ability and motivation (r = 0.47). Strong correlations were also found between these domains, between modulation of affects and reflective ability (r = 0.49), flexibility of interaction and reflective ability (r = 0.61), and flexibility of interaction and motivation (r = 0.43). Focus, the only measure in the nature of problems domain,

Table 11. Validity of the Suitability for Psychotherapy Scale (SPS): Mean values of the Quality of Object Relations (QORS, criterion measure) and the Symptom Checklist-90, Global Severity Index (SCL-90-GSI, discriminating measure) at baseline in categories of the seven individual suitability measures of the SPS (1-7) and the cumulative SPS score (0-7) (N = 326).

Suitability measure	Categories of the suitability measures								Correlation ratio	P-value for trend
	0	1	2	3	4	5	6	7		
A. Criterion validity	**Values of QORS**									
Nature of problems										
Focus		5.3	5.2	5.1	5.2	4.9	4.7	-	0.18	0.02
Ego strength										
Modulation of affects		5.5	5.3	5.1	4.9	5.1	4.2	-	0.30	< 0.001
Flexibility of interaction		5.7	5.2	5.1	4.7	4.9	4.7	4.5	0.39	< 0.001
Self-concept in relation to ego ideal		5.7	5.5	5.3	4.9	4.8	4.6	4.2	0.46	< 0.001
Self-observing capacity										
Response to trial interpretation		5.3	5.3	5.2	5.0	5.2	4.9	4.9	0.19	0.07
Reflective ability		5.5	5.3	5.0	4.9	4.7	4.6	-	0.38	< 0.001
Motivation		5.4	5.3	5.1	4.9	4.7	5.2	-	0.31	< 0.001
SPS score[1]	5.4	5.3	5.2	5.0	4.9	4.9	4.8	4.7	0.32	< 0.001
B. Discriminating validity	**Values of SCL-90-GSI**									
Nature of problems										
Focus		1.29	1.33	1.43	1.24	1.12	1.11	-	0.17	0.04
Ego strength										
Modulation of affects		1.14	1.30	1.27	1.26	1.41	1.29	-	0.12	0.20
Flexibility of interaction		1.27	1.28	1.22	1.45	1.15	1.55	1.16	0.15	0.70
Self-concept in relation to ego ideal		1.07	1.06	1.31	1.26	1.44	1.54	1.19	0.22	0.004
Self-observing capacity										
Response to trial interpretation		1.33	1.30	1.29	1.28	1.28	1.16	1.32	0.06	0.60
Reflective ability		1.35	1.27	1.26	1.16	1.42	1.55	-	0.13	0.44
Motivation		1.47	1.27	1.27	1.19	1.39	0.87	-	0.15	0.24
SPS score[1]	1.24	1.32	1.25	1.24	1.45	1.31	1.01	1.46	0.16	0.93

[1] The values of the seven individual suitability measures were dichotomized (good=0, poor=1) and summed up (0-7) (see Table 9).

Patient suitability for short-term
and long-term psychotherapy

Table 12. Correlations among the suitability measures: seven individual suitability measures of the Suitability for Psychotherapy (SPS) scale and the SPS score (N = 326).

Suitability measure	Categories of the suitability measures						
	1	2	3	4	5	6	7
Nature of problems							
1. Focus	1.00						
Ego strength							
2. Modulation of affects	0.25***	1.00					
3. Flexibility of interaction	0.31***	0.48***	1.00				
4. Self-concept	0.17***	0.19***	0.29***	1.00			
Self-observing capacity							
5. Trial interpretation	0.37***	0.28***	0.31***	0.15**	1.00		
6. Reflective ability	0.32***	0.49***	0.61***	0.32***	0.30***	1.00	
7. Motivation	0.30***	0.26***	0.43***	0.23***	0.19***	0.47***	1.00
8. *SPS score*	0.56***	0.50***	0.56***	0.35***	0.53***	0.56***	0.51***

* p-value < 0.05; ** p-value < 0.01; *** p-value < 0.001

correlated less strongly with the other suitability measures (r = 0.17-0.37). The SPS score correlated highly with all 7 individual suitability measures used in its calculation. The correlation coefficients for all other measures except self-concept (r = 0.35) were in the 0.50 range.

The suitability measures correlated statistically significantly only with the symptom outcome of short-term psychodynamic psychotherapy (STPP) at 1-year follow-up (Table 13). The smaller, and thus better, values in modulation of affects, flexibility of interaction with the interviewer, self-concept in relation to ego ideal, and reflective ability at baseline correlated with smaller, and thus better, values of the symptom measure SCL-90-GSI (r = 0.24-0.33). These correlations remained statistically significant at the 3-year follow-up point for modulation of affects (r = 0.23) and self-concept (r = 0.31). In addition, patient's good response to trial interpretation correlated statistically significantly with positive symptom outcome in STPP (r = 0.25). In solution-focused therapy (SFT), flexibility of interaction, reflective ability, and motivation were found to correlate statistically significantly with 3-year symptom outcome (r = 0.25-0.34). In long-term psychodynamic psychotherapy (LTPP) only reflective ability correlated with the symptom outcome at 3-year follow-up (r = 0.22). The cumulative SPS score correlated statistically significantly with the symptom outcome in STPP at 1-year follow-up (r = 0.30) and with both STPP and SFT symptom outcomes at 3-year follow-up (r = 0.29 and 0.28, respectively) but not with the LTPP symptom outcome.

Table 13. Correlations between the suitability measures and the psychotherapy outcome (SCL-90-GSI) at 12-month and 36-month follow-up in the three therapy groups.

Suitability measure	SCL-90-GSI					
	SFT		STPP		LTPP	
	12	36	12	36	12	36
Nature of problems						
Focus	0.06	0.12	0.20	0.08	0.09	0.05
Ego strength						
Modulation of affects	0.10	0.19	0.24[*]	0.23[*]	0.18	0.13
Flexibility of interaction	0.14	0.25[*]	0.32[**]	0.21	0.06	0.13
Self-concept	0.11	0.13	0.33[**]	0.31[**]	0.11	0.20
Self-observing capacity						
Trial interpretation	0.09	0.08	0.19	0.25[*]	0.06	0.03
Reflective ability	0.21	0.27[*]	0.29[**]	0.18	0.01	0.22[*]
Motivation	0.19	0.34[**]	0.05	0.06	0.05	0.01
SPS score	0.21	0.28[*]	0.30[**]	0.29[**]	0.09	0.11

[*] p-value < 0.05; [**] p-value < 0.01; [***] p-value < 0.001

5.3 Psychotherapy outcome prediction by the SPS (Studies II and III)

Prognoses for therapy outcome in STPP versus SFT (Study II) and short-term therapy (STPP and SFT combined) versus long-term therapy (LTPP) (Study III) were determined for patients assessed as having good and poor values in the seven individual suitability measures of the SPS and the cumulative SPS score, summing up the values of these seven measures.

5.3.1 Prediction by the SPS of the outcome of short-term psychotherapies (Study II)

In STPP, flexibility of interaction with the interviewer, self-concept, and reflective ability were the strongest individual predictors of psychiatric symptoms; patients with good values in these measures had a statistically significantly lower symptom level than patients with poor values throughout the 3-year follow-up, covering both the treatment (approximately 6 months) and the follow-up after the end of the treatment (approximately 2.5 years) (Table 14). Better elaboration in relation to trial interpretation and the modulation of affects differentiated the psychotherapy

outcome in the long run, during the follow-up after the end of the therapies (at 9, 12, 18, and 36 month follow-up points and 9, 18, and 36 month follow-up points, respectively). The focality of the complaints and level of motivation did not, however, differentiate the STPP outcome. The cumulative SPS score was a strong psychotherapy outcome predictor: on average, patients with mostly good values in the seven individual suitability measures (score values 0-3) experienced a 42% symptom reduction by the time of the 1-year follow-up in comparison to a 16% symptom reduction among patients with mostly poor values (score values 4-7); the first group also maintained the lowered symptom level in the 3-year follow-up, whereas the latter group did not.

In SFT, patients with a circumscribed focus had a significantly lower symptom level at the 7-month follow-up and patients with good reflective ability at the 7- and 9-month follow-up compared to patients with poor values in these measures (Table 14). A good cumulative SPS score predicted faster symptom reduction by the 9-month follow-up (data not shown). No other statistically significant differences were found between patients with good versus poor values in the individual suitability measures or the cumulative SPS score.

5.3.2 Comparison of the prediction by the SPS of the outcome of short-term psychotherapies (Study II)

A significant interaction between the level of motivation and form of short-term therapy (p = 0.04) and a nearly significant interaction between the level of self-concept in relation to ego ideal and form of short-term therapy (p = 0.06) was found (Table 14). Patients with good motivation and a realistic self-concept in relation to ego ideal seemed to benefit more from STPP, and patients with poor motivation and an unrealistic self-concept more from SFT. The only statistically significant difference in the symptom levels between STPP and SFT was, however, found among patients with an unrealistic self-concept both at the 7-month and 36-month follow-up points; in SFT these patients experienced a 36% symptom reduction on average in comparison to a 3% symptom reduction in STPP during the 3-year follow-up.

Patients with good values in the seven individual suitability measures and the cumulative SPS score experienced and maintained a statistically significant reduction in their psychiatric symptoms during the follow-up in comparison to the baseline symptom level both in STPP and in SFT (Table 14). The symptoms reduced 30%-45% during the first year of follow-up while the therapies were ongoing, and the attained symptom level was maintained during the second and third year of follow-up, the SCL-90-GSI varying between 0.71-0.87 at the 3-year follow-up point. No statistically significant differences in symptom prediction between STPP and SFT among those with good values in the suitability measures were found during the 3-year follow-up period.

Table 14. Mean values of the psychiatric symptoms (SCL-90-GSI) in short-term psychodynamic psychotherapy (STPP) and solution-focused therapy (SFT) and mean value differences (95% confidence intervals) between STPP and SFT at 0, 7, 12, 24, and 36 month follow-up according to the good and poor values of the seven individual suitability measures of the Suitability for Psychotherapy Scale (SPS) and the SPS score.

Suitability measure	Therapy	Good[1]					Poor[2]					P[3]
		0	7	12	24	36	0	7	12	24	36	
Nature of problems												
Focus	1. STPP	1.25	0.85	0.70	0.83	0.79	1.27	0.94	0.87	0.87	0.87	0.34
	2. SFT	1.28	**0.76**	0.82	0.81	0.80	1.33	0.99	0.90	0.98	0.88	
	1–2[4]		0.09	-0.12	0.02	-0.01		-0.05	-0.03	-0.11	-0.01	
			(-0.15,0.32)	(-0.36,0.12)	(-0.27,0.31)	(-0.28,0.31)		(-0.23,0.13)	(-0.21,0.16)	(-0.34,0.12)	(-0.22,0.21)	
Ego strength												
Modulation of affects	1. STPP	1.26	0.88	0.76	0.79	**0.73**	1.29	0.98	0.92	1.00	**1.04**	0.18
	2. SFT	1.28	0.89	0.89	0.91	0.87	1.36	0.94	0.82	0.91	0.82	
	1–2[4]		-0.01	-0.13	-0.13	-0.14		0.04	0.10	0.09	0.23	
			(-0.18,0.17)	(-0.31,0.06)	(-0.34,0.09)	(-0.35,0.07)		(-0.20,0.29)	(-0.16,0.35)	(-0.22,0.40)	(-0.07,0.52)	
Flexibility of interaction	1. STPP	1.28	**0.88**	**0.76**	**0.81**	**0.81**	1.17	**1.18**	**1.21**	**1.23**	**1.13**	0.97
	2. SFT	1.31	0.88	0.85	0.90	0.83	1.33	1.07	1.02	0.96	1.03	
	1–2[4]		-0.01	-0.09	-0.09	-0.02		0.11	0.19	0.27	0.11	
			(-0.16,0.14)	(-0.24,0.07)	(-0.28,0.10)	(-0.20,0.16)		(-0.31,0.53)	(-0.24,0.63)	(-0.27,0.81)	(-0.46,0.68)	
Self-concept in relation to ego ideal	1. STPP	1.26	**0.85**	**0.76**	**0.77**	**0.76**	1.31	**1.19**	1.09	**1.24**	**1.23**	0.06
	2. SFT	1.26	0.92	0.87	0.87	0.86	1.53	0.81	0.82	1.15	0.81	
	1–2[4]		-0.08	-0.12	-0.10	-0.10		0.39	0.27	0.09	0.43	
			(-0.23,0.08)	(-0.27,0.04)	(-0.30,0.10)	(-0.29,0.08)		(0.06,0.72)	(-0.06,0.60)	(-0.33,0.52)	(0.05,0.81)	
Self-observing capacity												
Response to trial interpretation	1. STPP	1.01	**0.85**	**0.74**	**0.78**	**0.71**	1.12	1.04	**0.96**	**1.03**	**1.08**	0.92
	2. SFT	0.98	**0.87**	0.84	0.90	0.83	1.09	0.97	0.96	0.95	0.95	
	1–2[4]		-0.02	-0.11	-0.12	-0.12		0.07	0.00	0.08	0.13	
			(-0.19,0.15)	(-0.28,0.07)	(-0.33,0.10)	(-0.32,0.08)		(-0.20,0.35)	(-0.27,0.27)	(-0.25,0.41)	(-0.19,0.45)	
Reflective ability	1. STPP	1.04	**0.86**	**0.76**	**0.80**	**0.79**	1.10	**1.13**	**1.06**	**1.14**	**1.10**	0.57
	2. SFT	1.00	**0.87**	0.84	0.91	0.84	1.02	**1.15**	1.00	0.90	0.91	
	1–2[4]		-0.00	-0.08	-0.11	-0.05		-0.02	0.07	0.24	0.19	
			(-0.16,0.15)	(-0.24,0.08)	(-0.30,0.09)	(-0.24,0.14)		(-0.35,0.31)	(-0.28,0.41)	(-0.19,0.67)	(-0.21,0.60)	

											P
Motivation											0.04
1. STPP	1.34	0.81	0.74	0.80	0.79	1.23	0.98	0.87	0.91	0.88	
2. SFT	1.18	0.92	0.89	0.94	0.78	1.39	0.89	0.85	0.88	0.88	
1–2[4]		-0.11	-0.15	-0.14	0.01		0.10	0.02	0.04	-0.00	
		(-0.34,0.12)	(-0.39,0.09)	(-0.42,0.14)	(-0.27,0.28)		(-0.08,0.28)	(-0.17,0.21)	(-0.20,0.27)	(-0.22,0.22)	
SPS score[5]											0.63
1. STPP	1.28	**0.86**	**0.74**	**0.78**	**0.74**	1.22	**1.12**	**1.08**	**1.15**	**1.17**	
2. SFT	1.27	0.86	0.84	0.86	0.84	1.45	1.08	0.99	1.06	0.93	
1–2[4]		-0.00	-0.10	-0.10	-0.09		0.05	0.10	0.09	0.24	
		(-0.16,0.16)	(-0.27,0.06)	(-0.29,0.10)	(-0.28,0.09)		(-0.27,0.36)	(-0.22,0.41)	(-0.30,0.48)	(-0.14,0.61)	

* The difference in the symptom values between STPP and SFT is statistically significant (P < 0.05); symptoms in bold differ statistically significantly (P < 0.05) between those with good values and those with poor values in the suitability measures; underlined symptoms have changed statistically significantly in comparison to baseline symptom level; symptoms in italic have decreased statistically significantly below suggested recovery level 0.9 on the SCL-90-GSI scale.

[1] Good values in the suitability measure. See Table 9 for classification of the suitability measures.

[2] Poor values in the suitability measure. See Table 9 for classification of the suitability measures.

[3] P for interaction between the suitability measure and the short-term therapy group throughout the follow-up.

[4] Mean value difference of SCL-90-GSI between STPP and SFT.

[5] The values of the seven individual dichotomous suitability measures (good=0, poor=1) were summed up (0-7). Values 0-3 of the cumulative score were considered good and values 4-7 poor.

Patients with a non-circumscribed focus, poor modulation of affects, and poor motivation also experienced and maintained a statistically significant symptom reduction during the follow-up in comparison to the baseline symptom level both in STPP and in SFT (Table 14). Their symptoms reduced up to 28%-35% during the first year of the follow-up, and the maintained average SCL-90-GSI scores varied between 0.82-1.04 at the 3-year follow-up point. Patients with poor flexibility of interaction with the interviewer, poor reflective ability, and poor response to trial interpretation did not experience and maintain a statistically significant symptom reduction either in STPP or SFT. Patients with an unrealistic self-concept in relation to ego ideal and patients with poor overall suitability did not experience a statistically significant symptom reduction in STPP but did experience and maintain such a symptom reduction in SFT. The only statistically significant difference in the symptom levels between STPP and SFT was, however, the above-mentioned difference among patients with an unrealistic self-concept.

5.3.3 Prediction by the SPS of the outcome of long-term psychotherapy (Study III)

In LTPP, patients with both good values and poor values in the seven individual suitability measures experienced and maintained a statistically significant symptom reduction during the 3-year follow-up in comparison to the baseline symptom level (Table 15); the symptom reduction was, however, greater among those with good values than among those with poor values in the suitability measures (46%-60% versus 31%-45%, respectively, at 3-year follow-up). No statistically significant differences in the psychiatric symptoms between the patients with good versus poor values in modulation of affects, flexibility of interaction with the interviewer, reflective ability, and motivation were found. Patients with better responsiveness to trial interpretation, on the other hand, experienced a faster symptom reduction in LTPP (data not shown), whereas patients with a more realistic self-concept and circumscribed focus experienced a greater symptom reduction in the long term, towards the end of long-term therapy (Table 15).

5.3.4 Comparison of the prediction by the SPS of the outcome of short-term versus long-term psychotherapies (Study III)

A significant interaction between circumscribed focus and length of therapy ($p = 0.05$) and a nearly significant interaction between the response to trial interpretation and length of therapy ($p = 0.08$) was found (Table 15). Symptoms of patients with either good or poor values in these two suitability measures seemed to reduce faster in short-term therapy, but more in long-term therapy in the long run. Only patients with a good response to trial interpretation, however, benefited statistically significantly faster from short-term than from long-term therapy, and only patients with a clearly circumscribed focus and poor response to trial

interpretation benefited statistically significantly more from long-term therapy at the 3-year follow-up point.

In addition to patients with a good response to trial interpretation, patients with good modulation of affects, flexibility of interaction, self-concept, and reflective ability also experienced a statistically significantly faster reduction in their symptoms in the short-term therapy group than in the long-term therapy group (36%-40% versus 26%-30%, respectively, during the first year of follow-up) (Table 15). In the long-term therapy group, where the treatment continued for approximately three years, the symptom level kept decreasing until the end of the 3-year follow-up, up to 47%-60%, whereas in the short-term therapy group the symptom level during the last two years of follow-up, when the treatment was no longer ongoing, remained at the level reached by the end of the first year of follow-up. Accordingly, by the 3-year follow-up point, in addition to the symptoms of patients with a clearly circumscribed focus, the symptoms of patients with good reflective ability had also decreased to a statistically significantly lower level in the long-term therapy group than in the short-term therapy group (SCL-90-GSI 0.51 versus 0.79, and 0.65 versus 0.79, respectively), and to a suggestively lower level among patients with good values in all other suitability measures.

Patients with poor values in the individual suitability measures, on the other hand, did not benefit statistically significantly more from short-term therapy than from long-term therapy at any point of the follow-up (Table 15). In the long-term therapy group, however, symptom reduction at the 3-year follow was statistically significantly larger than in short-term therapy group among patients with a poor self-concept, in addition to patients with a poor response to trial interpretation (SCL-90-GSI 0.70 versus 1.07, and 0.72 versus 1.03, respectively), and was suggestively larger among patients with poor values in all other suitability measures.

According to the cumulative SPS score, both patients with mainly good values in the seven individual suitability measures (score values 0-3) and patients with mainly poor values (score values 4-6) experienced and maintained a statistically significant symptom reduction both in the short-term and long-term therapy group, whereas patients for whom all seven suitability measures were poor (score value 7) did not (Table 16, Figure 2). Symptom reduction among those with mainly good values was faster in short-term therapy group than in long-term therapy group. Although patients with mainly poor values reached a much lower symptom level in the long-term than in short-term therapy group by the time of the 3-year follow-up (SCL-90-GSI 0.79 versus 1.03), this difference was not statistically significant. The difference between STPP and LTPP was, however, nearly significant (p = 0.05).

Overall, for patients who were assessed as having mostly good values in the suitability measures, short-term therapy seemed more beneficial due to its faster effect, whereas patients who were assessed as having mostly poor values seemed to benefit more from long-term therapy, and patients for whom all values were poor did not seem to benefit from either short-term or long-term therapy.

Table 15. Mean values of the psychiatric symptoms (SCL-90-GSI) in short-term and long-term therapy and mean value differences (95% confidence intervals) between short-term and long-term therapy at 0, 7, 12, 24, and 36 month follow-up according to the good and poor values of the seven individual suitability measures of the Suitability for Psychotherapy Scale (SPS).

Suitability measure	Therapy	Good[1]					Poor[2]					P[3]
		0	7	12	24	36	0	7	12	24	36	
Nature of problems												
Focus	1. Short	1.27	**0.80**	0.76	0.82	0.79	1.30	**0.97**	0.87	0.91	0.86	0.05
	2. Long	1.41	0.97	0.94	0.83	**0.51**	1.18	1.10	0.98	0.85	**0.78**	
	1-2[4]		-0.17	-0.18	-0.01	0.28		-0.14	-0.11	0.07	0.09	
			(-0.37,0.02)	(-0.36,0.01)	(-0.23,0.22)	(0.06,0.49)		(-0.29,0.02)	(-0.25,0.04)	(-0.11,0.24)	(-0.08,0.26)	
Ego strength												
Modulation of affects	1. Short	1.27	0.88	0.82	0.85	0.79	1.32	0.96	0.87	0.94	0.94	0.87
	2. Long	1.26	1.05	0.94	0.81	0.67	1.29	1.07	1.02	0.95	0.71	
	1-2[4]		-0.17	-0.12	0.04	0.12		-0.10	-0.16	-0.01	0.23	
			(-0.31,-0.03)	(-0.26,0.02)	(-0.12,0.20)	(-0.04,0.28)		(-0.33,0.13)	(-0.37,0.06)	(-0.26,0.24)	(-0.02,0.47)	
Flexibility of interaction[5]	1. Short	1.30	**0.88**	**0.79**	**0.85**	**0.80**	1.24	**1.12**	**1.12**	**1.16**	**1.15**	0.82
	2. Long	1.27	1.04	0.95	0.82	0.68	1.24	1.22	1.12	1.13	0.75	
	1-2[4]		-0.17	-0.15	0.03	0.12		-0.11	-0.00	0.03	0.40	
			(-0.29,-0.04)	(-0.27,-0.04)	(-0.11,0.17)	(-0.02,0.26)		(-0.57,0.36)	(-0.45,0.44)	(-0.43,0.49)	(-0.05,0.85)	
Self-concept in relation to ego ideal	1. Short	1.26	0.88	**0.80**	**0.82**	**0.78**	1.41	1.02	0.98	**1.17**	**1.07**	0.73
	2. Long	1.23	1.03	0.94	0.80	0.68	1.50	1.26	1.12	**1.11**	0.70	
	1-2[4]		-0.15	-0.14	0.02	0.10		-0.23	-0.14	0.05	0.37	
			(-0.28,-0.02)	(-0.26,-0.01)	(-0.13,0.16)	(-0.04,0.24)		(-0.56,0.10)	(-0.43,0.15)	(-0.28,0.39)	(0.05,0.69)	
Self-observing capacity												
Response to trial interpretation	1. Short	1.29	0.86	**0.79**	0.83	**0.75**	1.28	1.02	**0.94**	1.00	**1.03**	0.08
	2. Long	1.29	1.02	0.90	0.81	0.67	1.22	1.13	1.08	0.90	0.72	
	1-2[4]		-0.16	-0.11	0.02	0.08		-0.11	-0.14	0.11	0.30	
			(-0.30,-0.01)	(-0.25,0.03)	(-0.14,0.19)	(-0.08,0.24)		(-0.33,0.11)	(-0.34,0.06)	(-0.13,0.34)	(0.08,0.53)	
Reflective ability	1. Short	1.28	**0.86**	**0.80**	**0.84**	**0.79**	1.34	**1.14**	**1.03**	**1.07**	**1.05**	0.46
	2. Long	1.27	1.05	0.95	0.81	0.65	1.27	1.07	1.04	1.05	0.88	
	1-2[4]		-0.19	-0.16	0.03	0.15		0.06	-0.01	0.02	0.18	
			(-0.31,-0.06)	(-0.28,-0.03)	(-0.11,0.18)	(0.01,0.29)		(-0.24,0.37)	(-0.30,0.28)	(-0.33,0.35)	(-0.15,0.51)	

Motivation											
1. Short	1.26	0.85	0.81	0.86	0.79	1.31	0.94	0.85	0.89	0.87	0.63
2. Long	1.38	1.02	0.94	0.76	0.63	1.19	1.08	0.98	0.90	0.72	
1-2[4]		-0.17	-0.13	0.19	0.16		-0.14	-0.12	-0.01	0.15	
		(-0.36,0.02)	(-0.31,0.05)	(-0.11,0.31)	(-0.05,0.37)		(-0.30,0.02)	(-0.27,0.03)	(-0.18,0.17)	(-0.02,0.32)	

[*] The difference in the symptom values between short-term and long-term therapy is statistically significant (P < 0.05); symptoms in bold differ statistically significantly (P < 0.05) between those with good values and those with poor values in the suitability measures; underlined symptoms have changed statistically significantly in comparison to baseline symptom level; symptoms in italic have decreased statistically significantly below suggested recovery level 0.9 on the SCL-90-GSI scale.

[1] Good values in the individual suitability measure. See Table 9 for classification of the suitability measures.

[2] Poor values in the individual suitability measure. See Table 9 for classification of the suitability measures.

[3] P-value for interaction between the suitability measure and the therapy group throughout the follow-up.

[4] Mean value difference of SCL-90-GSI between short-term and long-term therapy.

[5] Simplified covariance structure.

Table 16. Mean values (95% confidence intervals) of the psychiatric symptoms (SCL-90-GSI) in short-term and long-term therapy at 0, 7, 12, 24, and 36 month follow-up according to the values of the Suitability for Psychotherapy Scale (SPS) score.

SPS score[1]	Short-term therapy						Long-term therapy						P^2
	N	0	7	12	24	36	N	0	7	12	24	36	
7 categories													0.70
0: All 7 good	27	1.18	0.74 (0.55,0.93)	0.70 (0.51,0.89)	0.78 (0.56,1.01)	0.67 (0.45,0.88)	13	1.38	0.89 (0.62,1.16)	0.89 (0.63,1.15)	0.76 (0.46,1.06)	0.61 (0.32,0.89)	
1: 6/7 good	41	1.27	0.88 (0.72,1.03)	0.93 (0.79,1.08)	0.90 (0.72,1.08)	0.85 (0.67,1.04)	31	1.40	1.11 (0.94,1.28)	0.95 (0.78,1.11)	0.81 (0.62,1.01)	0.67 (0.48,0.87)	
2: 5/7 good	47	1.31	0.94 (0.79,1.08)	0.70 (0.55,0.84)	0.80 (0.62,0.99)	0.73 (0.55,0.91)	41	1.18	0.91 (0.76,1.05)	0.87 (0.73,1.01)	0.74 (0.58,0.91)	0.61 (0.46,0.77)	
3: 4/7 good	39	1.32	0.85 (0.69,1.00)	0.76 (0.61,0.98)	0.80 (0.61,0.98)	0.84 (0.66,1.02)	17	1.07	1.21 (0.94,1.48)	1.01 (0.77,1.25)	0.82 (0.54,1.10)	0.78 (0.51,1.06)	
4: 3/7 good	17	1.46	1.02 (0.76,1.28)	0.98 (0.73,1.23)	0.98 (0.67,1.28)	0.95 (0.61,1.28)	10	1.43	1.06 (0.56,1.56)	1.30 (0.99,1.61)	0.94 (0.57,1.30)	0.61 (0.29,0.94)	
5: 2/7 good	13	1.24	1.05 (0.79,1.31)	0.85 (0.60,1.10)	1.03 (0.70,1.35)	1.03 (0.73,1.32)	9	1.42	1.31 (0.80,1.82)	1.05 (0.67,1.43)	1.40 (0.93,1.87)	0.72 (0.21,1.22)	
6: 1/7 good	8	1.07	1.19 (0.81,1.57)	1.19 (0.78,1.60)	1.32 (0.85,1.79)	1.16 (0.70,1.62)	5	0.91	1.27 (0.48,2.06)	1.21 (0.63,2.05)	1.19 (0.63,1.75)	0.84 (0.36,1.33)	
7: All 7poor	6	1.51	1.30 (0.86,1.73)	1.45 (1.05,1.84)	1.43 (0.99,1.88)	1.49 (0.98,2.00)	2	1.30	0.94 (0.05,1.82)	1.14 (0.32,1.97)	1.17 (0.19,2.14)	1.27 (0.33,2.20)	
3 categories													0.83
0-3	154	1.28	**0.86** (0.78,0.94)	**0.78** (0.70,0.86)	0.82 (0.72,0.91)	0.78 (0.69,0.87)	102	1.25	**1.02** (0.92,1.12)	**0.92** (0.83,1.03)	0.79 (0.68,0.90)	0.65 (0.54,0.76)	
4-6	38	1.30	1.05 (0.88,1.22)	0.95 (0.79,1.11)	1.05 (0.85,1.25)	1.03 (0.84,1.22)	24	1.31	1.23 (0.97,1.49)	1.18 (0.95,1.41)	1.12 (0.86,1.38)	0.79 (0.55,1.04)	
7	6	1.50	1.31 (0.87,1.76)	1.46 (1.06,1.86)	1.43 (1.00,1.87)	1.50 (1.00,1.99)	2	1.30	0.94 (0.01,1.88)	1.15 (0.28,2.02)	1.18 (0.15,2.20)	1.28 (0.27,2.28)	

Symptoms in bold differ statistically significantly (P < 0.05) between short-term and long-term therapy; underlined symptoms have changed statistically significantly in comparison to baseline symptom level; symptoms in italic have decreased statistically significantly below suggested recovery level 0.9 on the SCL-90-GSI scale.
[1] The values of the seven individual dichotomized suitability measures (good=0, poor=1) were summed up (0-7). Values 0-3 of the cumulative score were considered low, values 4-6 intermediate and value 7 high.
[2] P-value for interaction between the suitability measure and the therapy group throughout the follow-up.
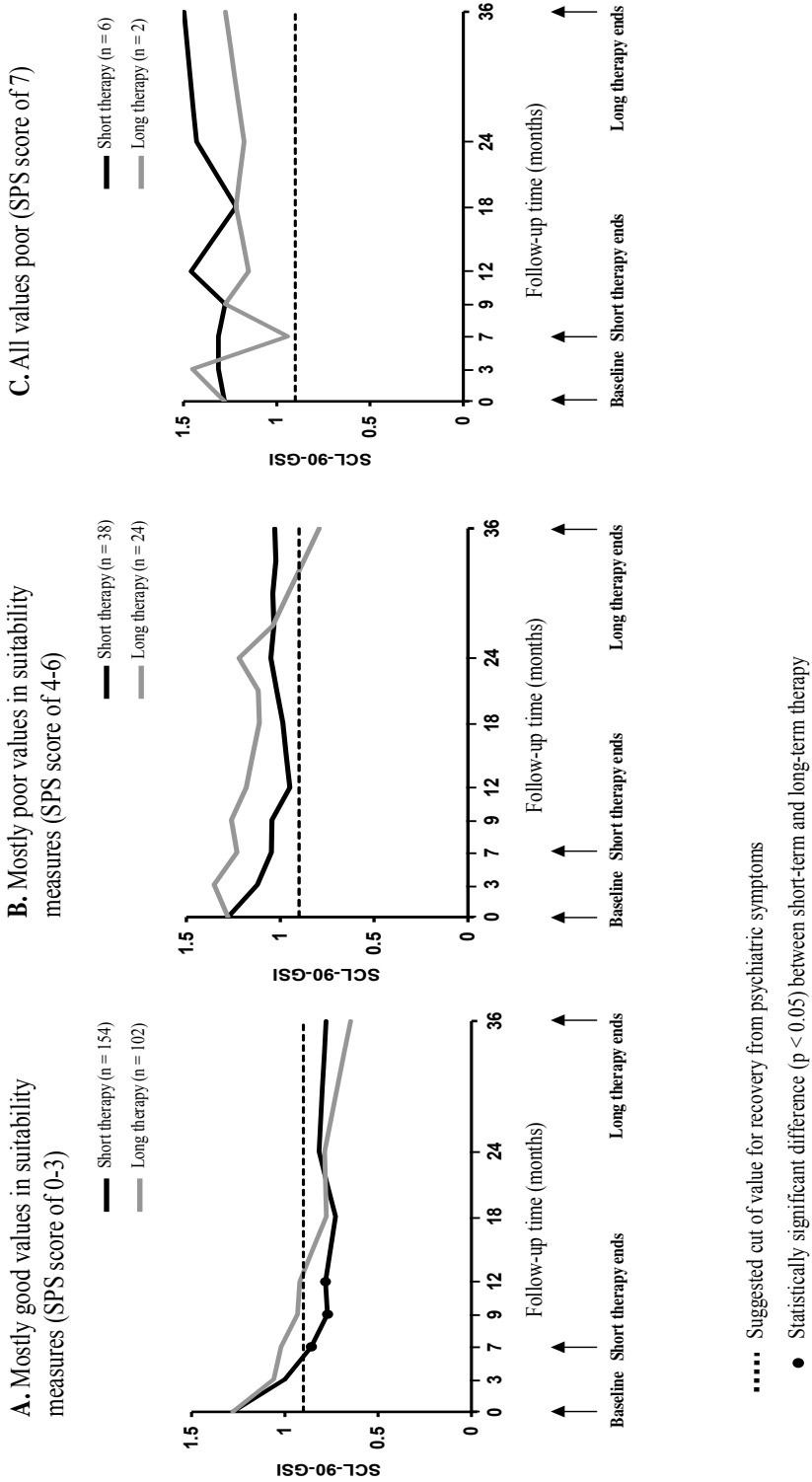
Figure 2. Changes in psychiatric symptoms in short-term and long-term therapy during a 3-year follow-up from the start of the treatments according to the SPS score.

Patient suitability for short-term
and long-term psychotherapy

6 DISCUSSION

This study showed that patient suitability for short-term and long-term psychotherapy can be predicted and differentiated by the new interview-based 7-item Suitability for Psychotherapy Scale (SPS) validated in this study.

6.1 Validation of the SPS (Study I)

6.1.1 Content and evaluation of the SPS

The SPS covered seven measures of a patient's psychological capacities suggested to predict and differentiate patient suitability for short-term versus long-term psychotherapy: circumscribed focus, modulation of affects, flexibility of interaction with the interviewer, self-concept in relation to ego ideal, response to trial interpretation, reflective ability, and motivation. It was hypothesized that good values in these suitability measures serve as indications for suitability for short-term psychotherapy, and poor values as contraindications for suitability for short-term psychotherapy but not for long-term psychotherapy (Malan 1976a, Sifneos 1972, Davanloo 1978, American Psychiatric Association 1985). Since good values in certain suitability measures may balance poor values in others, a cumulative SPS score, summing up the values of the seven individual suitability measures, was formed in order to better evaluate overall suitability. In addition to the seven suitability criteria included in the SPS, intelligence or problem-solving capacity, defense and coping styles, personality, and quality of interpersonal relationships have also been considered relevant in the evaluation of psychotherapy suitability. Well-validated and widely used measures for the evaluation of these suitability aspects, e.g., Wechsler Adult Intelligence Scale (WAIS; Wechsler 1955), Defense Style Questionnaire (DSQ; Andrews et al. 1989), and Quality of Object Relations Scale (QORS; Azim et al. 1991) were, however, already available at the time of the construction of the SPS, and were used for their evaluation in the Helsinki Psychotherapy Study (HPS). Patients' interpersonal capacities were also addressed via the SPS evaluation interview and assessment of the patients' flexibility of interaction with the interviewer, since clinical observation suggests that a patient's relationship style can be inferred in the very first interview from the way in which the patient interacts with the interviewer (Davanloo 1978). Prior to the SPS, 11 other psychotherapy suitability assessment scales or collections of suitability selection criteria had been presented in the literature (Heiberg 1975, Brodaty et al. 1982, Persson and Alström 1983, Buckley 1984, Piper et al. 1985, Vaslamatzis and Verveniotis 1985, Alpher et al. 1990, Safran et al. 1993, Rosenbaum et al. 1993, Fisher et al. 1999, Sigal et al. 1999). All except two scales (Safran et al. 1993, Fisher

et al. 1999) were aimed at evaluating suitability for interpretative short-term psychodynamic psychotherapy (STPP), within which the psychotherapy suitability criteria were first developed. Like the SPS, the majority of these other scales (Heiberg 1975, Brodaty et al. 1982, Piper et al. 1985, Vaslamatzis and Verveniotis 1985, Alpher et al. 1990, Safran et al. 1993, Rosenbaum et al. 1993) were relatively concise, consisting of 5-15 items; however, the largest scale (Fisher et al. 1999) comprised 226 items, making it relatively time-consuming to complete. The majority of the scales covered the evaluation of focus, modulation of affects, reflective ability, motivation, personality, and interpersonal relationships. For more than half of the scales an evaluation of overall suitability was also carried out (Heiberg 1975, Brodaty et al. 1982, Vaslamatzis end Verveniotis 1985, Alpher et al. 1990, Safran et al. 1993, Rosenbaum et al. 1993). However, unlike the SPS, which only included measures reflecting the patient's psychological capacities, several of the other scales (Persson and Alström 1983, Buckley 1984, Safran et al. 1993, Fisher et al. 1999, Sigal et al. 1999) also included other measures, such as measures of symptoms or working ability. This makes it difficult to separate the role of these measures from the suitability measures if such a scale is, for example, used to predict the outcome of psychotherapy.

The SPS was assessed based on three initial interview sessions with the aim of more reliably assessing the different aspects of the patient's suitability, as some patients are known to show inhibitions in the first interview, especially in communicating their inner emotional experiences, which are likely to be resolved by the second or third interview (Davanloo 1978). In the HPS, the assessment of the SPS was further integrated with the assessment of other potentially relevant aspects of suitability, such as the quality of interpersonal relationships, intelligence, and defense styles, and the diagnostic assessment (PDM Task Force 2006), which results in a comprehensive and efficient approach. The psychological suitability measures are considered rather complex phenomena that are difficult to assess, thus requiring knowledge of the content of these concepts and training for their evaluation. The interviews in the HPS were carried out by seven experienced interviewers – psychologists or psychiatrists with 9-20 years of experience in clinical practice and psychotherapy evaluation – who had received separate training for the interview procedure, including the evaluation of the SPS (60-100 hours). Both the reliability and the validity of the SPS assessments were thoroughly evaluated in this study. Most of the previously published suitability assessment scales were evaluated based on only one interview (Heiberg 1975, Persson and Alström 1983, Buckley 1984, Alpher et al. 1990, Safran et al. 1993, Fisher et al. 1999, Sigal et al. 1999). The experience and training of the interviewers for the suitability assessment was not reported in most of the published studies (Brodaty et al. 1982, Persson and Alström 1983, Buckley 1984, Piper et al. 1985, Vaslamatzis and Verveniotis 1985, Alpher et al. 1990, Sigal et al. 1999), but when it was reported (Heiberg 1975, Safran et al. 1993, Rosenbaum et al. 1997, Fisher et al. 1999), experienced assessors were used,

although the length of their training for the evaluation of the suitability assessment scale of interest varied from 10 hours (Fisher et al. 1999) to up to 5 years (Safran et al. 1993). Of the 11 previously published psychotherapy suitability assessment scales, 8 have been studied for reliability (Heiberg 1976, Barth et al. 1988a, b, Høglend et al. 1992a; Persson and Alström 1983; Buckley et al. 1984; Piper et al. 1985; Alpher et al. 1990, Baumann et al. 2001; Safran et al. 1993, Myhr et al. 2007; Rosenbaum et al. 1997; Fisher et al. 1999) and 5 also for validity (Heiberg 1976, Safran et al. 1993, Fisher et al. 1999, Jørgensen et al. 2000, Baumann et al. 2001).

6.1.2 Reliability of the SPS

In this study, the agreement between the seven individual interviewers' assessments of the seven individual SPS items, each rated on a 7-point scale, and the cumulative SPS score was found to be fair to good (median kappa coefficients 0.42-0.65) for all variables except for response to trial interpretation and motivation (median kappa coefficient 0.28 and 0.29, respectively). The largest variation in the agreement between interviewers was, however, also found for trial interpretation and motivation, with several interviewers reaching fair to good agreement. No systematic differences between the individual interviewers' assessments were found, suggesting that there were no relevant differences in the assessment techniques of different interviewers. The best agreement between individual interviewers was found for flexibility of interaction, modulation of affects, and reflective ability. This study was the first to report fair agreement on the evaluation of the patient's self-concept, and also the agreement on focality was found to be fair. Inter-rater reliability was measured based on a sample of 39 videotaped interviews. The agreement between interviewers was shown to be fair to good or excellent for all suitability items on the five previously evaluated psychotherapy suitability assessment scales (Heiberg 1976, Barth et al. 1988a, b; Safran et al. 1993, Myhr et al. 2007; Rosenbaum et al. 1997; Fisher et al. 1999; Baumann et al. 2001, Cromer and Hilsenroth 2010), vary greatly from poor to excellent for one scale (Persson and Alström 1983), and be poor to fair for all items on two scales (Høglend et al. 1992a, Piper et al. 1985). Half of these eight scales reported reliability separately for the individual suitability items (Persson and Alström 1983; Safran et al. 1993; Rosenbaum et al. 1997; Baumann et al. 2001, Cromer and Hilsenroth 2010). Similarly to this study, for these scales flexibility of interaction, modulation of affects, and reflective ability were also among the highest scoring suitability items (Safran et al. 1993; Rosenbaum et al. 1997; Baumann et al. 2001, Cromer and Hilsenroth 2010). The agreement on focality was also found to be fair in one previous study (Safran et al. 1993), although good agreement has also been reported (Persson and Alström 1983). Poor agreement on the aspects of motivation evaluated in this study, i.e., personal motivation for therapy and motivation for insight, has also been reported previously (Persson and Alström 1983). One potential reason for

poor agreement may be the interviewers' differential emphasis on the different aspects of motivation evaluated simultaneously. Agreement on trial interpretation was found to be good in the only previous study in which it was reported (Rosenbaum et al. 1997). However, Rosenbaum et al. (1997) evaluated the agreement on the patient's ability to consider clarification, confrontation, and interpretation, whereas this study evaluated only the patient's response to trial interpretation, which is likely to be more challenging and may thus explain the poorer agreement in this study. The inter-rater reliability in the previously published studies was measured based on samples of 10 to 167 interviews assessed by 2-7 interviewers (Heiberg 1976, Barth et al. 1988a, b, Høglend et al. 1992a; Persson and Alström 1983; Buckley et al. 1984; Piper et al. 1985; Alpher et al. 1990, Baumann et al. 2001, Cromer and Hilsenroth 2010; Safran et al. 1993, Myhr et al. 2007; Rosenbaum et al. 1997; Fisher et al. 1999). Samples larger than in this study (Heiberg 1976, Barth et al. 1988a, b; Cromer and Hilsenroth 2010; Rosenbaum et al. 1997; Fisher et al. 1999) were, however, assessed by 2-5 interviewers, typically only two, and the only interview sample evaluated by seven interviewers (Myhr et al. 2007) was smaller (N = 28) than in this study. In previous studies, the suitability items were typically rated on a 2- to 6-point scale, with the exception of two 9-point scales introduced by Alpher et al. (1990) and Safran et al. (1993), but later reduced to 5-point scales (Baumann et al. 2001, Cromer and Hilsenroth 2010; Myhr et al. 2007), one 8-point scale introduced by Høglend et al. (1992a), and one 7-point scale introduced by Piper et al. (1985). None of the previous studies analyzed whether systematic differences between the individual interviewers' assessments existed.

This was the first study to measure the repeatability of the interviewers' suitability assessments, making it possible to analyze the stability of the interviewers' assessment techniques over time. The repeatability of the individual interviewers' assessments three years apart, evaluated based on a subset of 27 videotaped interviews, was found to be nearly fair to good for all seven suitability measures. No notable systematic differences between the assessments made by the same interviewer over time were found in the present study, suggesting that there were no relevant differences in the assessment techniques of the same interviewers over time.

6.1.3 Validity of the SPS

In this study, the association between the patients' baseline scores of the SPS and the criterion measure QORS was strong, indicating good criterion validity. On the other hand, the association between the baseline scores of the SPS and the discriminating symptom measure SCL-90-GSI was weak, also suggesting good discriminating validity. Of the five suitability assessment scales previously studied for validity, three studied both criterion and discriminating validity (Heiberg 1976, Safran et al. 1993, Fisher et al. 1999), one only criterion validity (Jørgensen et al.

2000) and one only discriminating validity (Baumann et al. 2001). Two studies (Fisher et al. 1999, Jørgensen et al. 2000) assessed criterion validity and one study (Baumann et al. 2001) assessed discriminating validity by measuring the association between the suitability scale and an appropriate criterion or discriminating measure, in a similar manner to the present study. The rest of the studies, on the other hand, explored validity by measuring either the correlations between individual items of the suitability scale and an appropriate criterion or discriminating measure (Safran et al. 1993) or the intercorrelations between the individual items of the scale and comparing them to theory (Heiberg 1976, Fisher et al. 1999). The validity of all except one scale (DAI; Jørgensen et al. 2000) was supported.

In conclusion, the SPS appeared to be a reliable and valid method of assessing pre-treatment psychological suitability, and is the first validated interview-based psychotherapy suitability assessment scale to include the evaluation of the patient's self-concept and response to trial interpretation.

6.2 Psychotherapy outcome prediction by the SPS (Studies II and III)

6.2.1 Prediction by the SPS of the outcome of short-term psychotherapies (Study II)

This was the largest study (N = 101) on the prediction by the psychological suitability criteria of the outcome of interpretative short-term psychodynamic psychotherapy (STPP), for which the criteria were originally proposed. Of the 38 other datasets predicting symptom outcome of short-term psychotherapies by such suitability criteria, 11 included an interpretative STPP group comprising 18-72 patients. The majority of these 38 datasets focused, however, on the evaluation of the role of patients' intrapsychic and interpersonal behavior, especially the quality and quantity of patients' interpersonal relationships, in the outcome prediction. Less than half of the 38 datasets evaluated the role of at least one of the seven suitability criteria within the nature of problems, ego strength, and self-observing capacity suitability domains covered by the SPS, and 8 of these 18 datasets included STPP group (Brodaty et al. 1982; Horowitz et al. 1984; Husby et al. 1985, Husby 1985a; Høglend 1993; Piper et al. 1998; Sigal et al. 1999; Kronström et al. 2009; Cromer and Hilsenroth 2010). The other datasets studied cognitive-behavioral therapy (CBT; N = 5), interpersonal therapy (IPT; N = 4), behavioral therapy (BT; N = 2), and cognitive therapy (CT; N = 1). This study was the first to explore the adaptability of the suitability criteria to solution-focused therapy (SFT).

In STPP, 5 out of 7 individual SPS measures – modulation of affects, flexibility of interaction with the interviewer, self-concept in relation to ego ideal, response to trial interpretation, and reflective ability – and the SPS score, summing up the values

of the individual measures, predicted the symptom outcome. The majority of these suitability measures (flexibility of interaction, self-concept, reflective ability, SPS score) differentiated the symptom outcome between those with good versus poor values in these measures throughout the follow-up, i.e., both during the treatment and the follow-up after the treatment, but for some measures (trial interpretation, modulation of affects) these differences became statistically significant only during post-treatment follow-up, thus supporting the hypothesis of continued learning and improvement after the end of treatment (Sifneos 1972). In SFT, two SPS measures, focus and reflective ability, and the SPS score differentiated the symptom outcome, all predicting greater symptom improvement by the end of the therapy among those with better values for these measures.

Of the suitability measures reflecting ego strength, flexibility of interaction with the interviewer and self-concept in relation to ego ideal were among the strongest outcome predictors in STPP in this study, predicting a statistically significantly lower level of symptoms for patients with good values in these measures compared to patients with poor values throughout the 3-year follow-up, including an approximately 6-month treatment and a 2.5-year follow-up. Modulation of affects did not differentiate symptom improvement between the patients with good versus poor affect modulation during STPP treatment, but did differentiate this during the post-treatment follow-up: patients with good affect modulation reached a statistically significantly lower symptom level by the 3-year follow-up. In contrast to this study, a previous study on flexibility of interaction (Cromer and Hilsenroth 2010) and self-concept (Horowitz et al. 1984) did not find an association between their baseline level and psychiatric symptoms at the end of STPP treatment. Both studies measured the strength of the association using correlation, opposed to advanced modeling in this study, but additional correlation analysis in this study showed that also the correlations between these suitability measures and psychiatric symptoms were statistically significant. The sample sizes in the previous studies (71 and 51, respectively) were, however, smaller than in this study (N = 101). The results of two previous studies, of similar size and both studying modified versions of the Sifneos' selection criteria for Short-Term Anxiety-Provoking Psychotherapy (STAPP; Sifneos 1972, Heiberg 1975), with respect to associations between flexibility of interaction and psychiatric symptoms during the post-treatment follow-up were contradictory: one study found a strong association between adequate contact with the interviewer and psychiatric symptoms both at the 2-year (Husby et al. 1985) and at the 5-year (Husby 1985a) follow-up, in line with the findings from this study, whereas the other study found no association either at the 2-year or at the 4-year follow-up (Høglend 1993). In line with the findings of this study, no associations between modulation of affects and psychiatric symptoms at the end of STPP were found in three previous studies (Brodaty et al. 1982, Sigal et al. 1999, Cromer and Hilsenroth 2010), whereas a strong correlation between patients' ability to emote and psychiatric symptoms one year after the end of STPP was found in the

one study following patients beyond the end of treatment (Brodaty et al. 1982). In the only previously published study on associations between ego strength measures and symptom development in psychotherapy other than STPP, both awareness and differentiation of affects, and alliance potential, correlated strongly with psychiatric symptoms at the end of CBT (Myhr et al. 2007; N = 113).

Patients with better self-observing capacity, in terms of better reflective ability and response to trial interpretation, benefited more from STPP in this study. The baseline-adjusted symptoms of patients with better reflective ability were at a statistically significantly lower level throughout the 3-year follow-up, whereas patients with better elaboration on trial interpretation experienced a statistically significantly larger symptom reduction by the 1-year follow-up, which they maintained by the end of the 3-year follow-up. The level of motivation, on the other hand, did not predict the outcome of STPP. Regarding previous findings on reflective ability, also often referred to as psychological mindedness or insight, Piper et al. (1998; N = 64-66) found a statistically significant association between psychological mindedness and general psychiatric symptoms, but not with respect to depressive or anxiety symptoms, in interpretative, but not in supportive, STPP. This is in line with the findings of this study and also with theoretical suggestions that supportive STPP is also helpful to those not considered suitable candidates for the more anxiety-provoking interpretative STPP (Davanloo 1978, American Psychiatric Association 1985), with the level of suitability measures therefore separating the outcome of interpretative but not supportive STPP. Cromer and Hilsenroth (2010) found manifesting insight, but not appearing introspective or manifesting verbal fluency, to be associated with general symptoms and patient functioning at the end of STPP. One small study by Kronström et al. (2009; N =19) found no association between psychological mindedness and depressive symptoms either at the end of or eight months after interpretative STPP, whereas another small study (Brodaty et al. 1982; 18) did support an association between psychological mindedness and general symptoms one year after the end of STPP. The only previously published small study on the predictive role of response to trial interpretation in STPP (Brodaty et al. 1982; N = 18) found no association between patients' ability to accept interpretations and psychiatric symptoms either at the end of therapy or one year after the end of therapy. In line with the non-significant findings of this study, Brodaty et al. (1982), Horowitz et al. (1984) and Sigal et al. (1999) did not find any association between motivation and symptoms at the end of STPP, and neither did Brodaty et al. (1982) at the 1-year follow-up, Husby et al. (1985) at the 2-year follow-up, and Høglend (1993) at the 2- or 4-year follow-up after STPP. The only significant association was found by Husby (1985a) at the 5-year follow-up; five years is, however, a long follow-up time and the patients were not followed up for potential auxiliary treatment they had had by then. Association between some aspects of motivation and symptoms was supported in some other types of short-term therapies. In CBT, closely related motivational aspects related to acceptance of

personal responsibility for change (Myhr et al. 2007), autonomous, intrinsic, motivation to change (Zuroff et al. 2007), and readiness to change (Lewis et al. 2012) were found to be associated with a lower level of symptoms at the end of therapy, whereas patient expectation of improvement (Sotsky et al. 1991), optimism or pessimism regarding the therapy (Myhr et al. 2007), controlled motivation (Zuroff et al. 2007), and general motivation for therapy (Kampman et al. 2008) were not. Similarly to CBT, in cognitive therapy (CT) therapy readiness to change (Steketee et al. 2011), in BT the patient's willingness to participate (Keijsers et al. 1994a, b), and in interpersonal therapy (IPT) autonomous motivation (Zuroff et al. 2007, McBride et al. 2010), but not controlled motivation or expectation of improvement (Zuroff et al. 2007, Sotsky et al. 1991), were also associated with a lower level of symptoms at the end of the therapy.

Although a clearly circumscribed focus of treatment, agreed by both the patient and the therapist, has been considered theoretically highly important for suitability for STPP and short-term therapy in general (Balint et al. 1972, Davanloo 1978, Friedman and Lister 1987, Perry et al. 1987), no association between focus and psychiatric symptoms in STPP was found in this study or in any of the previously published studies (Brodaty et al. 1982; Husby et al. 1985, Husby 1985a, Høglend 1993; Sigal et al. 1999). Neither was any evidence found to support this association in CBT (Myhr et al. 2007). In this study, patients with a clearly circumscribed focus did, however, experience a greater symptom reduction by the end of SFT than patients with a poorly circumscribed focus.

The cumulative SPS score, summing up the values of the seven individual suitability measures of the SPS, was a strong predictor of the STPP outcome in this study; patients with good values in the majority of the suitability measures had a significantly lower level of symptoms throughout the 3-year follow-up than patients with mainly poor values in the suitability measures. In SFT, a good SPS score predicted greater symptom reduction by the end of therapy. The previously published results on the STPP outcome prediction by the summary scores of two suitability scales supported their prediction (Brodaty et al. 1982, Alpher et al. 1990) and thus the findings of this study. Alpher et al. (1990) found a better total score on the validated Capacity for Dynamic Process Scale (CDPS) to be associated with a lower level of general symptoms at the end of therapy in a relative small sample of 25 patients; a later, larger study of 71 patients by Cromer and Hilsenroth (2010) was, however, unable to confirm this finding on the CDPS total score. The sum of the six individual suitability items of the non-validated selection criteria studied by Brodaty et al. (1982) did not correlate with the symptom outcome at the end of STPP but did correlate with the symptom outcome one year after the therapy. In CBT, high overall suitability according to the validated Suitability for Short-term Cognitive Therapy (SSCT) Rating Scale was found to predict lower symptoms at the end of the therapy in both published studies (Safran et al. 1993, Myhr et al. 2007), even after adjustment for the baseline symptom level and confounding factors.

In conclusion, this study supported the ability of the suggested criteria to differentiate between suitability for STPP, with those assessed as having good values in the suitability measures generally benefiting more from STPP than those assessed as having poor values. The individual SPS measures and the SPS score were stronger predictors of the STPP outcome than the SFT outcome where they rarely differentiated between those who were likely to benefit more and less.

6.2.2 Comparison of the prediction by the SPS of the outcome of short-term psychotherapies (Study II)

Identifying the measures that predict the outcome of different therapy modalities helps to indicate which kinds of patients seem especially suitable for them. To know, however, which of those treatments suits the patient best, requires knowledge on whether patients with certain resources benefit more from one therapy than from another. This requires comparison of the prediction by the suitability measures between different psychotherapy modalities.

This study was the first to compare the prediction by the psychological psychotherapy suitability criteria of psychiatric symptoms in STPP and another type of short-term therapy, i.e., SFT. These suitability criteria were thus adapted to other types of short-term therapies without accumulating similar empirical evidence from clinical practice on patients treated with respective therapy as that gathered with respect to STPP – based on which most of these criteria were first suggested – or comparing whether prediction by the suitability measures empirically supported in psychodynamic psychotherapy was similar in other types of short-term therapies. Of the 38 datasets evaluating suitability for short-term therapy, only five included more than one type of short-term therapy, and only three of these evaluated the same suitability measures as the SPS. One dataset compared prediction by reflective ability in interpretative and supportive forms of STPP (Piper et al. 1998), whereas the two other datasets compared prediction by motivation in CBT and IPT (Sotsky et al. 1991, Zuroff et al. 2007).

In this study, motivation appeared to be a significant, and self-concept in relation to ego ideal a nearly significant, predictor of differential symptom outcome in STPP versus SFT: patients with good motivation and a realistic self-concept seemed to benefit more from STPP and patients with poor motivation and an unrealistic self-concept more from SFT. No statistically significant differences in the symptom levels between STPP and SFT among patients with good values in motivation or self-concept, or any of the other five individual suitability measures or the SPS score were, however, found at any point in the follow-up (Study II). Neither were any statistically significant differences in symptom levels found between the two short-term therapies among patients with poor values in the suitability measures, with the exception of patients with an unrealistic self-concept, whose symptoms reduced faster in SFT than in STPP and were still lower at the 3-year follow-up point. One

explanation for this may lie in the goal-focused and resource-oriented nature of SFT, which was also reflected in a significantly more positive therapist-rated working alliance (Knekt and Lindfors 2004), which may facilitate better symptomatic change for patients characterized by self-denigration than the transference-focused interpretative technique applied in STPP. Similarly to this study, no significant interaction between reflective ability and the form of short-term therapy was found by Piper et al. (1998): better reflective ability was significantly associated with a larger symptom reduction in interpretative STPP, in line with findings of this study, but not in supportive STPP, in line with theoretical suggestions that supportive STPP is also beneficial for patients with poorer psychological capacities. Unlike in this study, no significant interaction between motivation and the form of short-term therapy was found by Sotsky et al. (1991) and Zuroff et al. (2007), who compared the prediction by motivation in CBT versus IPT; autonomous motivation was found to be a predictor of a good symptom outcome in both (Zuroff et al. 2007) and the expectation of improvement in neither (Sotsky et al. 1991).

In conclusion, patients with good values in the SPS measures had an equally good prognosis in STPP and SFT. Patients with poor values in the SPS measures also had an equal prognosis in STPP and SFT, except for patients with an unrealistic self-concept, who benefited more from SFT. Thus, although the differences between the symptom levels of patients assessed as having good versus poor values in the SPS measures were more pronounced in STPP than in SFT, as mentioned above, those assessed as having equal resources, good or poor, did not differ in their short-term psychotherapy outcome prediction.

6.2.3 Prediction by the SPS of the outcome of long-term psychotherapy (Study III)

This was the first study to evaluate prediction by circumscribed focus, flexibility of interaction with the interviewer, self-concept in relation to ego ideal, response to trial interpretation, reflective ability, and motivation of the symptom outcome of LTPP. Only three other studies, all based on separate datasets, have so far evaluated the association between psychological suitability measures and symptom outcome in long-term therapies (Jørgensen et al. 2000, Puschner et al. 2004, Solbakken et al. 2012). All three datasets included LTPP and one dataset (Puschner et al. 2004) also included long-term cognitive therapy (LTCT). Of the 13 suggested psychological suitability selection criteria (see Table 1), all of which have been studied at least once in order to determine their prediction in short-term therapies, only modulation of affects (Solbakken et al. 2012), quality of interpersonal relationships (Jørgensen et al. 2000, Puschner et al. 2004), and overall suitability (Jørgensen et al. 2000) have been studied in respect to long-term therapies. Suitability research has thus so far almost entirely focused on short-term therapies. One apparent reason for this is the

laborious nature of long-term therapy studies: they are expensive and time-consuming. Furthermore, prediction by the suitability measures is likely to be stronger, and thus to differentiate the patients' therapy outcome better, in short-term therapies than in long-term therapies, where both patients with good and poor values in the suitability measures are theoretically expected to benefit from treatment. Of course, it may be hypothesized that patients with good values in the suitability measures benefit more from long-term therapy than patients with poor values, and thus significant differences between such patient groups may be observed, but a greater likelihood of significant findings in short-term therapies may also have played a role in directing the focus of suitability research. This study (N = 128) was not the largest LTPP suitability study so far – both the study by Puschner et al. (2004; N = 397) and the study by Solbakken et al. (2012; N = 153) were larger – but the average length of LTPP in this study (232 sessions) was much larger than the average length of LTPPs in previous studies (around 70 sessions). This study thus represents better long-term psychodynamic psychotherapy, whereas the previous studies have rather represented medium-term psychodynamic psychotherapy.

In this study, patients with both good and poor values in the SPS measures experienced and maintained a statistically significant symptom reduction in LTPP, and no statistically significant differences in the psychiatric symptoms between the patients with good versus poor values were found in the majority of the suitability measures (modulation of affects, flexibility of interaction with the interviewer, reflective ability, and motivation) at any point during the follow-up. This supports the hypothesis that LTPP also benefits patients with poor values in the suitability measures, and suggests that LTPP may even be equally beneficial for patients with good and poor values in some of the suitability measures. Patients with good values in the suitability measures may, however, also be in a better position compared to patients with poor values in long-term therapy, as demonstrated by patients with better responsiveness to trial interpretation benefiting faster from LTPP and patients with a more realistic self-concept and circumscribed focus gaining more from LTPP in the long term. Patients with better overall suitability, as measured by the SPS score, benefited more from LTPP throughout the follow-up. The only previously published, smaller study on the role of overall suitability in LTPP (Jørgensen et al. 2000; N = 16) did not find that the overall suitability assessed based on the Dynamic Assessment Interview (DAI) differentiated the LTPP outcome. Whereas in this study the level of modulation of affects did not differentiate the LTPP outcome, in a previous study by Solbakken et al. (2012) patients with more severe impairment in affect consciousness and integration were found to benefit even more from LTPP. Similarly, both patients making positive and negative interpersonal contributions experienced significant changes in psychiatric symptoms in LTPP (Jørgensen et al. 2000), and patients demonstrating hostility or hostile submission were found to experience the greatest symptom reduction in LTPP (Puschner et al. 2004). Findings therefore exist which support equal treatability with LTPP of patients with good or

poor values in certain suitability measures, better treatability of patients with good values in certain suitability measures, and better treatability of patients with poor values in certain suitability measures. As none of these findings have so far been replicated, more research on the predictive role of these suitability measures in LTPP is needed in order to confirm these findings. More research is also needed on the suitability measures which have not yet been studied in LTPP (i.e., intelligence, coping styles and cognitive skills, and personality measures) and on all suitability measures in other types of long-term therapies. Furthermore, only one study (Puschner et al. 2004) has so far compared the prediction by psychological suitability measures of the outcome of two different forms of individual long-term therapies, namely LTPP and LTCT. Whereas LTPP, intended to last 50-80 sessions, was relatively short in this German study and was thus considered to represent medium-term rather than long-term therapy, LTCT, intended to last 25-50 sessions, was shorter than the generally considered minimum length of a long-term therapy of at least 40-50 sessions (Leichsenring and Rabung 2008, de Maat et al. 2009, Smit et al. 2012). No studies comparing the prediction by suitability measures between two truly long-term therapies thus currently exists, and more studies on the comparison between different forms of long-term therapies are urgently needed.

In conclusion, both patients with good and poor values in the studied suitability measures benefited from the LTPP. Patients with mainly good values in the suitability measures benefited more from LTPP than patients with mainly poor values. This study thus supported vast treatability by LTPP.

6.2.4 Comparison of the prediction by the SPS of the outcome of short-term versus long-term psychotherapies (Study III)

This was the first study to compare the prediction by the psychological suitability measures of the outcome of individual short-term versus long-term psychotherapy, although the psychotherapy suitability criteria were particularly developed for differentiating patients suitable for and treatable with short-term psychotherapy from those in need of long-term psychotherapy (Davanloo 1978). Patients not fulfilling the criteria for the short-term psychotherapy of interest, included in the previously published and validated psychotherapy suitability assessment scales, were thus apparently assumed to be suitable for long-term psychotherapy, but this was neither explicitly stated nor studied. However, although patients with certain contraindications, i.e., poor values in the suitability measures, were shown not to benefit from short-term psychotherapy, they may not benefit from long-term psychotherapy either, in which case the effectiveness of some other treatments should be studied. Furthermore, although patients with certain indications, i.e., good values in the suitability measures, were shown to benefit from short-term psychotherapy, they may still benefit more from long-term psychotherapy, in which

case long-term psychotherapy should, if available and affordable, still be the recommended treatment option. Confirming that the suggested psychotherapy suitability selection criteria actually work in practice thus requires evidence that a) short-term psychotherapy is more beneficial than or as beneficial as long-term psychotherapy in the treatment of patients with good values in the suitability measures, and that b) long-term psychotherapy only is beneficial or more beneficial than short-term psychotherapy in the treatment of patients with poor values in the suitability measures. Only then can these measures truly be applied in the selection of patients for short-term versus long-term psychotherapy.

In this study, patients with good values in the suitability measures experienced an enduring statistically significant symptom reduction both in short-term and long-term psychotherapy. However, the symptoms of patients with good values in the modulation of affects, flexibility of interaction, self-concept, response to trial interpretation, and reflective ability reduced at a statistically significantly faster rate in short-term than in long-term psychotherapy. In the long run, however, the symptoms of patients with a clearly circumscribed focus and good reflective ability reduced statistically significantly more in the long-term psychotherapy. Short-term psychotherapy was thus clearly more beneficial than or as beneficial as long-term psychotherapy among patients with good values in modulation of affects, flexibility of interaction, self-concept, response to trial interpretation, and motivation, whereas long-term psychotherapy was more beneficial among patients with a clearly circumscribed focus. Patients with good reflective ability improved faster in short-term psychotherapy but gained more from long-term psychotherapy in the long run; this makes the appraisal of which therapy was more beneficial more difficult but in most cases the long-term outcome is given more importance. According to the SPS score, patients with mainly good values in the seven individual suitability measures (i.e., with an SPS score of 0-3) experienced faster symptom reduction in short-term psychotherapy than in long-term psychotherapy and did not benefit more from long-term psychotherapy at any point of the 3-year follow-up. The SPS score could thus identify those patients who were best treatable with short-term psychotherapy.

Patients with poor values in the suitability measures also experienced an enduring statistically significant symptom reduction both in short-term and long-term psychotherapy, except for patients with poor flexibility of interaction with the interviewer whose symptoms reduced statistically significantly only in long-term psychotherapy. The average symptom reduction was, however, much larger in long-term psychotherapy than in short-term psychotherapy; it was statistically significantly larger among patients with a poor self-concept and response to trial interpretation and suggestively larger among patients with poor values in all other suitability measures. Long-term psychotherapy was thus beneficial for all patients with poor values in the individual suitability measures and was evidently more beneficial than short-term psychotherapy among patients with poor values in flexibility of interaction, self-concept, and response to trial interpretation.

Furthermore, among patients with mainly poor values in the seven individual suitability measures (i.e., with an SPS score of 4-6) the difference in the symptom levels at the 3-year follow-up point between STPP and LTPP was nearly significant. Use of the SPS score also identified a group of patients with poor values in all seven suitability measures (i.e., with an SPS score of 7) who did not seem to benefit from short-term or long-term therapy as their symptoms remained at the same level as at the start of the therapy, or even increased during the 3-year follow-up. Such patients apparently need treatment other than the kinds offered in this study, but due to the very small sample size of eight patients in this subgroup, a chance finding cannot be excluded either. More research is thus needed in order to verify that lack of several psychological capacities is a contraindication not only for short-term psychotherapy, but also for long-term psychotherapy.

In conclusion, this study was the first to support differential prediction by the psychological suitability measures of the outcome of short-term versus long-term psychotherapy, and thus the functionality of the psychotherapy suitability selection criteria as originally intended. Self-concept and response to trial interpretation were prescriptive variables, predicting faster symptom improvement in short-term therapy among patients with good values in relation to such variables and greater symptom improvement in long-term therapy among patients with poor values in relation to such variables. Modulation of affects, flexibility of interaction with the interviewer, and reflective ability were prognostic variables predicting faster symptom improvement in short-term therapy among patients with good values in relation to such variables, and focality was a prognostic variable predicting larger symptom improvement in long-term therapy among patients with a circumscribed focus. The SPS score predicted different patterns of outcome between short-term and long-term therapy, with faster symptom reduction in short-term therapy among patients with mainly good values in the seven suitability measures, and suggestively larger symptom reduction in long-term therapy among patients with mainly poor values.

6.3 Methodological considerations

6.3.1 Validation of the SPS (Study I)

The new Suitability for Psychotherapy Scale (SPS) and its evaluation in the present study (Study I) had several strengths but also some weaknesses. First, the SPS assessed the level of seven different aspects of patients' psychological capacity, considered relevant, either as indications or contraindications, for suitability for short-term and long-term psychotherapy, but largely lacking validated measures for their evaluation. The three previously published psychotherapy suitability assessment scales (STAPP, CDPS, and SSCT), shown to be reliable, valid, and predictive of symptom outcome in short-term psychotherapies, cover many of the

same suitability dimensions as the SPS, but none of them evaluate patients' self-concept and patients' response to trial interpretation, which were the two strongest outcome predictors in this study. Other aspects, such as patients' intelligence, defense style, and interpersonal relations have also been considered relevant for psychotherapy suitability, but were not included in the SPS since other well-validated and widely used measures for their evaluation already existed. Second, a cumulative SPS score, capturing simultaneously the seven individual aspects, was formed to allow the evaluation of the overall importance of these capacities, some of which the same individual may master better than others and which may balance each other. Such summary scores have been shown to be best to predict psychotherapy outcome, yet the majority of the suitability studies have concentrated on individual capacities. The summary score was calculated by summing up the values of the dichotomized, comparable individual suitability measures, and was further categorized for the final analysis. Although such categorized suitability measures are helpful in the description and comparison of the results, categorization may cause loss of information. As some of the suitability measures were also more correlated than others, such an approach may emphasize some dimensions of suitability over the others in the overall evaluation. An alternative approach would be to use factorization to first group the individual suitability measures and then form an overall score based on the factor scores. More research on optimal ways of summarizing the suitability information is needed. Third, the assessment of the SPS was based on an interview, which is considered a more reliable and objective method for assessing the suitability measures reflecting the patient's psychological capacities, which may be partly unconscious and possibly affected by the psychiatric disorder, than the patient's self-evaluation via a questionnaire (Fisher et al. 1999, Rosenbaum et al. 1997). Moreover, two of the suitability criteria included in the SPS, patient's flexibility of interaction and response to trial interpretation, can only be evaluated based on interaction between the patient and the interviewer. The semi-structured interview procedure used, in which the relevant information is sought out via direct interaction with the interviewer, is considered to reflect the same dynamics as psychotherapy itself, and can thus be considered well founded for the purpose of the assessment of suitability (Rosenbaum et al. 1997). The evaluation of the SPS was carried out during three pre-therapy interview sessions to allow more profound evaluation of suitability, as some aspects, such as expression of affects, response to trial interpretation, and motivation, may not yet emerge during the first interview. Fourth, because of the complex nature of psychological suitability measures, only experienced interviewers thoroughly trained for the method were chosen for the present study. A comprehensive picture of reliability was guaranteed by using seven different interviewers. Fifth, the SPS, evaluated by such interviewers, was shown to be a reliable and valid method for assessing patient's psychotherapy suitability. The repeatability of the individual interviewers' suitability assessments over time, reported for the first time in this study, was fair or

good for all SPS items. The agreement between the individual interviewers' assessments was also fair or good for all SPS items except for response to trial interpretation and motivation. Assessment of these suitability aspects thus warrants special attention when training for the method in future to guarantee their uniform assessment. Both criterion and discriminating validity of the SPS were good, indicating that the scale truly measures what it was intended to measure.

6.3.2 Psychotherapy outcome prediction by the SPS (Studies II and III)

The design and implementation of the studies on the prediction by the suitability measures, evaluated using the SPS, of the outcome of short-term and long-term psychotherapies (Studies II and III) had several advantages. First, these studies were based on data from the HPS which is the first randomized clinical trial comparing the effectiveness of individual short-term and long-term psychotherapies (Knekt and Lindfors 2004, Knekt et al. 2008a), and thus also the first study allowing the evaluation of suitability for short-term and long-term psychotherapy in a comparative design. This was thus the first study to compare the prediction by the psychological suitability measures of the outcome of individual short-term and long-term psychotherapy, including both short-term and long-term psychodynamic psychotherapies (STPP and LTPP) for which the suitability selection criteria were first proposed. This was also the first study to compare the prediction by these criteria of the outcome of short-term psychodynamic psychotherapy and another short-term therapy, i.e., solution-focused therapy (SFT), and thus to evaluate their adaptability to SFT, and on the whole the first study on the prediction by these criteria of the outcome of SFT. Optimally, to study the differential prediction by the suitability measures, patients would be randomized to the therapies based on their values in those measures, to allow an even distribution of patients with good and poor values in the suitability measures and with factors potentially confounding the relationship between the suitability measures and psychotherapy outcome in the different therapy groups. Mere randomization of the patients to the therapies in the HPS was, however, shown to result in fairly equal distribution of patients with good and poor values in the suitability measures in different therapy groups (Studies II and III). Thorough statistical analyses and adjustment for confounding factors were carried out to further ensure fair comparison of the psychotherapy outcome prediction by these suitability measures. Second, the sample size of this study was among the largest suitability study sample sizes, and the largest STPP sample size, therefore increasing the possibility to detect relevant effects. Since the two short-term therapies were shown to not differ from one another either in their suitability (Study II) or effectiveness (Knekt et al. 2008), they could be combined into one short-term therapy group, thus further increasing the sample size in the short-term therapy group. The LTPP provided in this study was the longest of the few suitability studies on LTPP, and therefore better representative of long-term rather

than medium-term psychotherapy. Psychotherapy research, particularly on long-term therapies, is expensive and time-consuming, and a large sample of patients that have been treated long-term is thus difficult to obtain. The large sample size of this study allowed the interaction analysis between the suitability measures and therapy groups but due to the small sample size of some of the subgroups further larger studies on them are especially needed to confirm the findings. Third, all the therapists were qualified and experienced in the therapy orientations they provided (Knekt et al. 2008a). Fourth, the long follow-up time with frequently repeated outcome measurements offered the possibility to estimate and compare the patients' symptom profiles for the different treatments (Knekt et al. 2008a); these symptom profiles were very different for short-term versus long-term treatments which has previously largely been ignored. Fifth, the participation rate was high throughout the follow-up, resulting in comprehensive repeated measurement data. Additional analyses based on data completed by multiple imputation were, however, also performed to analyze the potential effect of drop-out (Härkänen et al. 2005). No notable differences between the analyses based on the original and the imputed data were found, however. Sixth, both the evaluation of suitability, using the SPS (Study I), and psychotherapy outcome, using the SCL-90-GSI, were based on validated methods. Seventh, the baseline symptom level in the treatment groups compared was adjusted by including the symptom variable at baseline as a covariate in the model, to prevent possible bias due to different baseline levels of the outcome variable. Also other potential confounding factors (i.e., socio-demographic factors, psychiatric history, and personality functions) were thoroughly adjusted for, to ensure that the results were due to the independent prediction by the suitability measures. The confounding factors were also evaluated using validated methods. Eight, a fundamental advantage of this study was that for the first time comprehensive data on auxiliary treatment (psychotropic medication, non-protocol psychotherapy, or psychiatric hospitalization), which would have been unethical to deny and reflects patients' continuing need for treatment, were collected during the entire follow-up, and this data collection also covered patients refusing the treatment offered (Knekt et al. 2011b). This made it possible to estimate the recovery from psychiatric symptoms both ignoring – as has so far been done in psychotherapy studies – and considering the effect of auxiliary treatment, thus allowing better evaluation of the independent effect of the study treatment. Thus, due to thorough evaluation of the confounding factors and auxiliary treatment, the data could be analyzed using both intention to-treat and as-treated designs, with basic and complete models, to explore other factors possibly affecting the results in addition to the suitability measures of interest. Ninth, advanced modeling based on linear mixed models and inference based on standard measures of strength of association, adjusted mean differences and their confidence intervals, were used. Tenth, new more comprehensive criteria for the evaluation of the prediction by the suitability measures of psychotherapy outcome, assessing statistical significance of different

dimensions of change, were presented and applied. These criteria included the evaluation of statistical significance of both change in outcome in time and difference in outcome between patients with good and poor values of the suitability measures and between therapy groups.

Several issues may, however, complicate the interpretation of the results. First, because of the long follow-up, no non-treatment control group could be included in the HPS design for ethical reasons (Knekt and Lindfors 2004, Knekt et al. 2008a). Consequently, the possible reduction in symptoms due to factors other than the treatments given, such as regression to the mean, could not be controlled. Since the symptom change in the two short-term therapies during the first few months was quite similar (Study II) but different from the symptom change in long-term therapy (Study III), it appeared that the treatment length had a clear effect on symptoms, and accordingly regression to the mean did not seem to affect the comparability between the treatment groups. Second, some of the patients had to wait much longer for their treatment to start than others which also might have biased the comparisons (Knekt et al. 2008a), especially if waiting times were unevenly distributed between the different therapy groups, as it is well known that the symptoms of the patients are at their highest when seeking treatment. Adjustment for the waiting times in the statistical models did not, however, change the results, which suggests than uneven waiting times did not bias the results. Third, since the aim of the HPS was to study the effectiveness of and suitability for psychotherapy given for outpatients in normal clinical practice, no treatment manuals were used in the psychodynamic psychotherapies (Knekt et al. 2008a). A larger number of therapists, especially in LTPP covering various theoretical models but also in STPP, were used to ensure the generalizability of the results. Fourth, although potential confounding factors were adjusted by modeling, the possibility of residual confounding due to unmeasured confounding factors cannot be fully excluded. Fifth, the compliance with the study treatment (i.e., withdrawal from the treatment after randomization, discontinuation of the treatment, and use of auxiliary treatment during the follow-up) may potentially have caused bias in the data. One fifth of the patients in the LTPP group withdrew from treatment after randomization and one fifth of those starting the treatment discontinued prematurely; these patients may have experienced weaker symptoms and might therefore have been unwilling to commit themselves for three years (Knekt et al. 2008a). In short-term therapies only 3%-4% refused the treatment and 10%-12% discontinued it but the patients discontinuing SFT had more symptoms than those completing the treatment; these patients may have felt too sick to continue the treatment or may have thought that the treatment was not helping them (Knekt et al. 2008a). The occurrence of auxiliary treatment was lower in the LTPP group during the 3-year follow-up than in the short-term treatment groups; the patients in short-term therapies may have felt greater need for treatment during or after the short-term therapies and the use of auxiliary treatment may have been reflected in the results in addition to the effects of the study treatments (Knekt et al.

2011b). Adjustments for withdrawal, discontinuation, and auxiliary treatment in as-treated analysis did not, however, notably alter the results from those of the intention-to-treat analysis. Sixth, although the drop-out rate during the follow-up was low, those who dropped out from the SFT more often had psychiatric symptoms and more often needed psychiatric treatment, which may have biased the results in the primary analyses based on the original data and the assumption of ignorable drop-outs (Knekt et al. 2008a). No notable differences were, however, found between the intention-to-treat and as-treated analysis, carried out based on both the original and imputed data, which suggests the results to be unbiased. However, despite the relatively frequent repeated measurements, especially during the early follow-up, some of the questions were formulated to describe the patient's current status at that point of follow-up (e.g., current use of psychotropic medication), and thus some relevant information between the measurement points, especially towards the end of the follow-up, may have been lost (e.g., starting and stopping psychotropic medication between two measurement points up to one year apart from each other). Advanced modeling (Härkänen et al. 2013) and further research with different or additional assessment can help to clarify the importance of such issues. Seventh, the follow-up time in this study was three years from the start of the treatments, during which the patients were provided with either short-term therapy, followed by no treatment, or long-term therapy, to allow the comparison of the prediction by the suitability measures of outcome of short-term versus long-term therapy. Due to the shorter duration of short-term therapies, this study also provides information on the stability of the findings in short-term therapies, but not in long-term therapies, in the long run. The symptom outcome at the end of short-term therapies was on average quite stable. According to the 5-year HPS effectiveness results (Knekt et al. 2013), the symptom level reached at the 3-year follow-up point in the long-term psychodynamic psychotherapy was maintained during the next two years, which indirectly implies that the conclusions of this study regarding the long-term therapy are also likely to hold in a longer follow-up. Eight, due to the multiple comparisons performed the possibility of a chance finding cannot be excluded and due to the relatively small amount of data in certain subgroups, such as patients with poor values in the suitability measures in long-term therapy group (i.e., SPS score values 4-7) and patients with all suitability measures poor (i.e., the highest value of the SPS score), a relevant difference may not have been detected.

6.4 Conclusions and implications for further research

In conclusion, although criteria for patient suitability for psychotherapy were developed half a century ago, prior research on their reliable evaluation and ability to predict and differentiate the outcome of short-term and long-term psychotherapy was scarce. This study showed that the new Suitability for Psychotherapy Scale (SPS) can be reliably applied to the assessment of the seven suitability measures

included in it, and is apparently the first reliable and valid psychotherapy suitability assessment scale to include the assessment of the patient's self-concept and response to trial interpretation. These two suitability measures best differentiated the outcome of short-term versus long-term psychotherapy in the study: patients with a realistic self-concept and good response to trial interpretation experienced faster symptom reduction in short-term psychotherapy, whereas patients with an unrealistic self-concept and poor response to trial interpretation experienced greater symptom reduction in long-term psychotherapy. More generally, patients assessed as having mainly good values in the seven suitability measures studied could apparently be successfully treated with short-term therapy, whereas patients assessed as having mainly poor values seemed to benefit more from long-term therapy. This was thus the first study to provide evidence of the ability of the suggested psychotherapy suitability selection criteria to differentiate those suitable for short-term therapy from those in need of long-term therapy. Suitability for the two different types of short-term therapy did not, on the other hand, seem to differ notably. The choice between therapy length thus seems more relevant than that between therapy type, and the amount of therapy needed can apparently be predicted before the start of therapy using a compact psychotherapy suitability assessment scale such as the SPS.

6.4.1 Replication and generalization of the findings

More research is needed to verify the findings of this study and to investigate whether they also extend to other types of psychotherapies. More research is also needed in order to identify whether the findings regarding self-reported psychiatric symptoms can be extended to other symptom measures, both self-reported and observer-rated, and to other outcome measures, such as work ability, and social and psychological functioning. Various well-validated outcome measures, both self-reported and observer-rated, were evaluated in the HPS (Knekt and Lindfors 2004), and two other suitability studies supported the similar prediction of both self-reported and observer-rated symptoms (Laaksonen et al. 2014), and of work ability and social and psychological functioning (Lindfors et al. 2012, Laaksonen et al. 2014). This supports a successful generalization of the findings of this suitability study across other symptom and other outcome measures. Given that the psychotherapy experienced is useful, the greatest changes in patients' well-being are likely to occur and be observed in the outcome measure reflecting the reasons for seeking psychotherapy. In the HPS, an enquiry was made about the patient's reasons for seeking psychotherapy, using the Target Complaints scale (Battle et al. 1966). Further study can therefore also be conducted on whether the use of the outcome measure best reflecting patient's target complaints further enhances psychotherapy outcome prediction by the SPS.

In addition to the seven suitability criteria included in the SPS and covered by the present study, the prediction by the other suggested suitability criteria, i.e.,

intelligence, defense and coping styles, personality, and quality of interpersonal relationships, of the outcome of short-term versus long-term psychotherapy should also be studied. In the HPS, intelligence, defense style, and the quality of interpersonal relationships were also evaluated (Knekt and Lindfors 2004, Knekt et al. 2010), and findings published so far on the prediction by defense style (Laaksonen et al. 2014) and quality of interpersonal relationships (Lindfors et al. 2014) resemble the findings of this study. The SPS could thus potentially be combined with these other suitability measures in order to form a comprehensive psychotherapy suitability selection tool and score. Besides psychological suitability, other patient factors, such as psychiatric symptoms and diagnosis (Knekt et al. 2008a) and socio-economic factors (Joutsenniemi et al. 2012), have also been shown to differentiate the outcome of short-term and long-term psychotherapy in the HPS. Knowledge of the mutual importance of the different patient factors remains fragmentary, however, and requires further study. After identifying the best independent predictors, their relative importance to psychotherapy outcome prediction and differentiation can be determined using statistical measures such as the Population Attributable Fraction (PAF), which assesses the fraction of the outcome attributable to different predictors (Laaksonen 2010). Such information could be used in forming an even more comprehensive, evidence-based screening instrument for psychotherapies of different length.

The study population in this study comprised patients suffering from mood and anxiety disorders, which represent the largest groups of psychiatric disorders and are generally considered treatable with short-term psychotherapy. The scope of patients considered treatable with short-term psychotherapy has, however, continually expanded, and nowadays patients with more severe, complex, and treatment resistant disorders, previously considered to require longer treatments, have also been suggested to be potentially treatable with short-term therapies (Abbass et al. 2011, Solbakken and Abbass 2014). Further research is needed to clarify whether psychotherapy suitability selection criteria could also differentiate the treatability of such patients with short-term versus long-term therapy.

In the replication of the findings derived from the suitability studies and their expansion to cover other therapy forms, outcome measures, suitability measures, and psychiatric disorders, use of validated, publicly available suitability and outcome measures should be strongly promoted. Such studies should also be based on sufficiently large patient samples, chosen based on generally agreed and clearly stated inclusion and exclusion criteria, evaluated and treated by experienced interviewers and therapists, followed up beyond the end of treatment, and analyzed using consistent methods of a higher standard, in order to guarantee the detection, replication, and comparison of the relevant results. The replication and generalization of the suitability findings may, however, raise ethical concerns since patients assessed as having some poor values in the suitability measures may not be expected to benefit from short-term therapy, at least optimally, and patients with

only poor values in the suitability measures may not even be expected to benefit from long-term therapy. On the other hand, for a certain treatment to be judged effective and research findings to be considered valid, the evidence-based criteria (American Psychiatric Association 1993, American Psychological Association 1995, GRADE Working Group et al. 2004) require that several high-quality studies – preferably randomized clinical trials – replicating the findings have been carried out. Discussion of ethical considerations related to the application of evidence-based criteria to suitability research is thus needed.

6.4.2 Explanation of the findings

The mechanisms via which the suitability profiles predict the psychotherapy outcome are unknown. They may be related to the original status of the suitability factors, or possible changes in them during the therapy process. As the psychological capacities have been considered relatively stable, the apparently greater improvement in the patients with mainly poor values in such suitability measures in longer treatment may be explained by the notion that only longer treatment is capable of improving these capacities. SPS measures have been repeatedly assessed during short-term and long-term treatments in a small HPS subsample, thus providing the possibility to study and compare the potential changes in them in order to evaluate the correctness of this hypothesis. Patients with a lack of such capacities may also pose more challenges to the therapist and to the development of the patient-therapist alliance, which is also known to influence the psychotherapy outcome, thus requiring more time in order for the treatment to be effective. Poorer psychological capacities do not, however, self-evidently result in poorer alliance, and a poor outcome prognosis may very well also be compensated, or a good outcome prognosis further enhanced, by good alliance. More research on the factors explaining or modifying the predicted suitability profiles is thus needed.

The stability of the psychotherapy outcome also warrants more research. The findings of this study on the persistence, or even further improvement, of the symptom level achieved by the end of short-term therapies in further follow-up are in accordance with the theoretical suggestion that the learning process initiated during therapy continues beyond the end of therapy and possibly brings patients further benefits (Sifneos 1972). However, HPS patients, particularly those assigned to short-term therapies, also often sought auxiliary treatment after the end of the protocol therapy, thus indicating a continuing need for treatment, and possibly partly explaining the stability of the outcome findings. Very few previous suitability studies have followed patients beyond the end of their treatment, even fewer have reported auxiliary treatment, and none have taken auxiliary treatment into account in their analysis. More research is thus needed in order to explore the potential association between the stability of the outcome findings and the incidence of auxiliary treatment, as well as other factors potentially explaining the stability

identified. As some HPS patients sought auxiliary treatment after the end of their short-term therapy and some patients ended their long-term therapy prematurely, it may very well be that there is a patient group in need of medium-term, rather than short-term or long-term therapy.

6.4.3 Application of the findings

Ultimately, the verified psychotherapy suitability selection criteria should be applied in a real-world clinical patient population in order to demonstrate their value in practice. This could be done, for example, by assessing a certain number of patients from an ongoing patient stream using the psychotherapy suitability assessment instrument, and then referring half of the patients for treatment according to current clinical practice and the other half according to the new suitability-based referral instrument. The referred patients would then be followed throughout and beyond their treatment, in order to compare whether the treatment referral based on the new instrument results in a better outcome, which is persistent over time. Such data could also be used for studying the potential reasons for inaccurate outcome prognosis and non-response, in order to further improve the outcome prediction. The development of a well-functioning psychotherapy referral tool which better informs the current clinical practice can be expected to have a strong impact on decision-making practices and to enhance more rational and cost-effective use of health care resources.

7 ACKNOWLEDGEMENTS

This study was carried out as part of the Helsinki Psychotherapy Study (HPS) at the National Institute for Health and Welfare (THL). The HPS is the world-first randomized clinical trial on the effectiveness of short-term and long-term psychotherapies in the treatment of mood and anxiety disorders, initiated in the early 1990s in collaboration between THL, Biomedicum Helsinki, the Social Insurance Institution, Finland, the Psychiatric Clinic of Helsinki University Central Hospital, the Rehabilitation Foundation, and the Finnish Student Health Service. I feel truly privileged to have been able to be part of this study for the last ten years and to be among the first researchers to study and report patient suitability for short-term and long-term therapy based on such rigorous research data. I would like to thank all the patients, the therapists, and the members of the HPS group whose contribution and commitment to the HPS made the collection of this remarkable data source and implementation of this thesis possible. I also gratefully acknowledge the financial support by Kela, the Academy of Finland (Grant nos. 138876 and 253088), the Emil Aaltonen Foundation, and the Foundation for Psychiatric Research in Finland both to the HPS and to me, and the excellent research facilities provided by THL.

My greatest appreciation and gratitude belongs to my supervisors who have been my main source of support and guidance throughout my doctoral studies. First and foremost I would like to thank my primary supervisor and the director of the HPS, Professor Emeritus Paul Knekt, PHD. He offered me this novel topic for my research and guided me through every stage of my study, teaching me all of the concepts and methods I needed to know and brainstorming with me to generate new methods and approaches needed for this study. His passion and dedication to research, enormous amount of knowledge and expertise in several areas, and unrivaled ability to plan, coordinate, and lead cutting-edge research has been a true inspiration to me. Despite his multiple projects and countless work commitments, his door was always open to me and he was always ready to share his knowledge and advice. I have been greatly privileged to have him as my supervisor in my two Master's theses and doctoral dissertations and I can only hope to be able to follow his exceptional example with my own students. I also warmly thank my secondary supervisor Doctor Olavi Lindfors, PhD, whose extensive clinical knowledge and insight has been invaluable to my research. His long experience both as a psychoanalyst and a psychotherapy researcher, as well as his involvement in psychotherapy training and evaluation, both in the HPS and in the broader psychotherapy community, has provided him with a unique set of skills and expertise which have been extremely useful and helpful for carrying out the various psychotherapy suitability classifications and interpretations in this study. He has also been responsible for the collection of the HPS data. I am also very thankful to

8 REFERENCES

Abbass A, Town J, Driessen E (2011). The efficacy of short-term psychodynamic psychotherapy for depressive disorders with comorbid personality disorder. *Psychiatry*, 74, 58-71.

Abraham RE, Van HL, van Foeken I, Ingenhoven TJ, Tremonti W, Pieper-de Vries I, de Groot J, van 't Spijker A, Spinhoven P (2001). The developmental profile. *J Pers Disord*, 15, 457-473.

Alden LE, Wiggins JS, Pincus AL (1990). Construction of circumplex scales for the Inventory of Interpersonal Problems. *J Pers Assess*, 55, 521-536.

Alexander F (1925). Review of Die Entwicklungsziele der Psychoanalyse by S. Ferenczi and O. Rank. *Internationale Zeitschrift für Psychoanalyse*, 11.

Alexander F, French TM (1946). *Psychoanalytic therapy*. New York, NY: Ronald Press.

Alexander FG (1944). The Indications for Psychoanalytic Therapy. *Bull N Y Acad Med*, 20, 319-332.

Alonso J, Angermeyer MC, Bernert S, Bruffaerts R, Brugha TS, Bryson H, de Girolamo G, Graaf R, Demyttenaere K, Gasquet I, Haro JM, Katz SJ, Kessler RC, Kovess V, Lepine JP, Ormel J, Polidori G, Russo LJ, Vilagut G, Almansa J, Arbabzadeh-Bouchez S, Autonell J, Bernal M, Buist-Bouwman MA, Codony M, Domingo-Salvany A, Ferrer M, Joo SS, Martinez-Alonso M, Matschinger H, Mazzi F, Morgan Z, Morosini P, Palacin C, Romera B, Taub N, Vollebergh WA (2004). Prevalence of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. *Acta Psychiatr Scand Suppl*, 420, 21-27.

Alpher VS, Henry WP, Strupp HH (1990). Dynamic factors in patient assessment and prediction of change in short-term dynamic psychotherapy. *Psychother*, 27, 350-361.

American Psychiatric Association (1985). *Manual of psychiatric peer review* (3rd ed.). Washington, DC: American Psychiatric Association.

American Psychiatric Association (1993). Practice guidelines for major depressive disorder in adults. *Am J Psychiatry*, 150, 1-26.

American Psychiatric Association (1994). *Diagnostic and statistical manual of mental disorders: DSM-IV* (4th ed.). Washington, DC: American Psychiatric Association.

American Psychological Association (1995). Task Force Promotion Disseminating Psychological Procedures, Division of Clinical Psychology. Training in and dissemination of empirically-validated psychological treatments: report and recommendations. *Clin Psychol*, 48, 3-23.

Ammons RB, Ammons CH (1962). *Quick Test*. Missoula, MT: University of Montana.

Andrews G, Pollock C, Stewart G (1989). The determination of defense style by questionnaire. *Arch Gen Psychiatry*, 46, 455-460.

Andrews G, Singh M, Bond M (1993). The Defense Style Questionnaire. *J Nerv Ment Dis*, 181, 246-256.

Antonovsky A (1993). The structure and properties of the sense of coherence scale. *Soc Sci Med*, 36, 725-733.

Appelbaum A (1972). A critical re-examination of the concept 'motivation for change' in psychoanalytic treatment. *Int J Psychoanal*, 53, 51-59.

Azim HF, Piper WE, Segal PM, Nixon GW, Duncan SC (1991). The Quality of Object Relations Scale. *Bull Menninger Clin*, 55, 323-343.

Baardseth TP, Goldberg SB, Pace BT, Wislocki AP, Frost ND, Siddiqui JR, Lindemann AM, Kivlighan DM 3rd, Laska KM, Del Re AC, Minami T, Wampold BE (2013). Cognitive-behavioral therapy versus other therapies: redux. *Clin Psychol Rev*, 33, 395-405.

Bachrach HM, Galatzer-Levy R, Skolnikoff A, Waldron S Jr. (1991). On the efficacy of psychoanalysis. *J Am Psychoanal Assoc*, 39, 871-916.

Baker E (1947). Brief psychotherapy. *J Med Soc N J*, 44, 260.

Balint M, Ornstein PH, Balint E (1972). *Focal psychotherapy. An example of applied psychoanalysis*. London: Tavistock Publications.

Barron F (1953). An ego-strength scale which predicts response to psychotherapy. *J Consult Psychol*, 17, 327-333.

Barth K, Havik OE, Nielsen G, Haver B, Molstad E, Rogge H, Skatun M, Noklebye Heiberg A, Ursin H (1988b). Factor analysis of the evaluation form for selecting patients for short-term anxiety-provoking psychotherapy. The Bergen project on brief dynamic psychotherapy. *Psychother Psychosom*, 49, 47-52.

Barth K, Nielsen G, Havik OE, Haver B, Molstad E, Rogge H, Skatun M, Heiberg AN, Ursin H (1988a). Assessment for three different forms of short-term dynamic psychotherapy. Findings from the Bergen Project. *Psychother Psychosom*, 49, 153-159.

Battle CC, Imber SD, Hoehn-Saric R, Nash ER, Frank JD (1966). Target complaints as criteria of improvement. *Am J Psychother*, 20, 184-192.

Baumann BD, Hilsenroth MJ, Ackerman SJ, Baity MR, Smith CL, Smith SR, Blagys MD, Price JL, Heindselman TL, Mount MK, Holdwick DJ (2001). The capacity for dynamic process scale, an examination of reliability, validity, and relation to therapeutic outcome. *Psychother Res*, 11, 275-294.

Beck AT (1970). *Depression: Causes and treatment*. Philadelphia, PA: University of Pennsylvania Press.

Beck AT (1976). *Cognitive therapy and the emotional disorders*. New York, NY: Meridian.

Beck AT, Rush AJ, Shaw BF, Emery G (1979). *Cognitive therapy of depression*. New York, NY: Guildford Press.

Beck AT, Ward CH, Mendelson M, Mock J, Erbaugh J (1961). An inventory for measuring depression. *Arch Gen Psychiatry*, 4, 561-571.

Beck AT, Weissman A, Lester D, Trexler L (1974). The measurement of pessimism: The Hopelessness Scale. *J Consult Clin Psychol*, 42, 861-865.

Bellak L, Small L (1965). *Emergency psychotherapy and brief psychotherapy*. New York, NY: Grune and Stratton.

Benjamin LS (1974). Structural analysis of social behavior. *Psychological Review*, 81, 392-425.

Benjamin LS (1983). *The INTREX user's manual (Pts. I and II)*. Madison, WI: INTREX Interpersonal Institute.

Benjamin LS (1996). A clinician-friendly version of the interpersonal circumplex: structural analysis of social behavior (SASB). *J Pers Assess*, 66, 248-266.

Berghout CC, Zevalkink J, Katzko MW, de Jong JT (2012). Changes in symptoms and interpersonal problems during the first 2 years of long-term psychoanalytic psychotherapy and psychoanalysis. *Psychol Psychother*, 85, 203-219.

Berliner B (1941). Short psychoanalytic psychotherapy: its possibilities and its limitations. *Bull Menninger Clin*, 5, 204-213.

Beutel ME, Leuzinger-Bohleber M, Ruger B, Bahrke U, Negele A, Haselbacher A, Fiedler G, Keller W, Hautzinger M (2012). Psychoanalytic and cognitive-behavior therapy of chronic depression: study protocol for a randomized controlled trial. *Trials*, 13, 117.

Bland JM, Altman DG (1994). Regression towards the mean. *BMJ*, 308, 1499.

Blatt SJ, D'Afflitti JP, Quinlan DM (1976). Experiences of depression in normal young adults. *J Abnorm Psychol*, 85, 383-389.

Blatt SJ, Quinlan DM, Pilkonis PA, Shea MT (1995). Impact of perfectionism and need for approval on the brief treatment of depression: the National Institute of Mental Health Treatment of Depression Collaborative Research Program revisited. *J Consult Clin Psychol*, 63, 125-132.

Blenkiron P (1999). Who is suitable for cognitive behavioural therapy? *J R Soc Med*, 92, 222-229.

Blom MB, Jonker K, Dusseldorp E, Spinhoven P, Hoencamp E, Haffmans J, van Dyck R (2007a). Combinationt treatment for acute depression is superior only when psychotherapy is added to medication. *Psychother Psychosom*, 76, 289-297.

Blom MB, Spinhoven P, Hoffman T, Jonker K, Hoencamp E, Haffmans PM, van Dyck R (2007b). Severity and duration of depression, not personality factors, predict short term outcome in the treatment of major depression. *J Affect Disord*, 104, 119-126.

Bond M, Gardner ST, Christian J, Sigal JJ (1983). Empirical study of self-rated defense styles. *Arch Gen Psychiatry*, 40, 333-338.

Bond M, Perry JC, Gautier M, Goldenberg M, Oppenheimer J, Simand J (1989). Validating the self-report of defense styles. *J Pers Disord*, 52, 289-301.

Bordin ES (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychother Theory Res Pract*, 16, 252-260.

Braun SR, Gregor B, Tran US (2013). Comparing bona fide psychotherapies of depression in adults with two meta-analytical approaches. *PLoS One*, 8, e68135.

Brenner H, Kliebsch U (1996). Dependence on weighted kappa coefficients on the number of categories. *Epidemiology*, 7, 199-202.

Breuer J, Freud S (1895). Studies on Hysteria. In: J. Strachey (Ed.). *The standard edition of the complete psychological works of Sigmund Freud*. London: Hogarth Press, 1955, 2.

Brodaty H, Andrews G, Grant WB (1982). An attempt to predict who will benefit from brief psychotherapy in a general practice setting. *Aust N Z J Psychiatry*, 16, 69-73.

Buckley P, Conte HR, Plutchik R, Wild KV, Karasu TB (1984). Psychodynamic variables as predictors of psychotherapy outcome. *Am J Psychiatry*, 141, 742-748.

Budd R, Hughes I (2009). The Dodo Bird Verdict-- controversial, inevitable and important: a commentary on 30 years of meta-analyses. *Clin Psychol Psychother*, 16, 510-522.

Burdon AP (1963). Principles of brief psychotherapy. *J La State Med Soc*, 115, 374-378.

Buss AH, Durkee A (1957). An inventory for assessing different kinds of hostility. *J Consult Psychol*, 21, 343-349.

Caligor E, Hilsenroth MJ, Devlin M, Rutherford BR, Terry M, Roose SP (2012). Will patients accept randomization to psychoanalysis? A feasibility study. *J Am Psychoanal Assoc*, 60, 337-360.

Carter JD, Luty SE, McKenzie JM, Mulder RT, Frampton CM, Joyce PR (2011). Patient predictors of response to cognitive behaviour therapy and interpersonal psychotherapy in a randomised clinical trial for depression. *J Affect Disord*, 128, 252-261.

Chambless DL, Hollon SD (1998). Defining empirically supported therapies. *J Consult Clin Psychol*, 66, 7-18.

Churchill R, Hunot V, Corney R, Knapp M, McGuire H, Tylee A, Wessely S (2001). A systematic review of controlled trials of the effectiveness and cost-effectiveness of brief psychological treatments for depression. *Health Technol Assess*, 5, 1-173.

Clark LA (1993). *Schedule for Nonadaptive and Adaptive Personality (SNAP)*. Minneapolis, MN: University of Minnesota Press.

Clark LA, Vittengl J, Kraft D, Jarrett RB (2003a). Separate personality traits from states to predict depression. *J Pers Disord*, 17, 152-172.

Clark LA, Vittengl JR, Kraft D, Jarrett RB (2003b). Shared, not unique, components of personality and psychosocial functioning predict depression severity after acute-phase cognitive therapy. *J Pers Disord*, 17, 406-430.

Cloninger CR, Pryzbeck TR, Svrakic DM, Wetzel RD (1994). *The Temperament and Character Inventory (TCI): A guide to its development and use*. St Louis, MO: Center for Psychobiology of Personality, Washington University.

Cohen J (1960). A coefficient of agreement for nominal scales. *Edu and Psych Meas*, 20, 37-46.

Cohen J (1968). Weighted kappa: nominal scale agreement with provision for scaled disagreement or partical credit. *Psych Bull*, 70, 213-220.

Cohen J, Cohen P (1975). *Applied multiple regression / correlation analysis for the behavioural sciences*. New York, NY: Wiley.

Conte HR, Plutchik R, Jung BB, Picard S, Karasu TB, Lotterman A (1990). Psychological mindedness as a predictor of psychotherapy outcome: a preliminary report. *Compr Psychiatry*, 31, 426-431.

Cooper P, Osborn M, Gath D, Feggetter G (1982). Evaluation of a modified self-report measure of social adjustment. *Br J Psychiatry*, 141, 68-75.

Costa PT, Mccrae RR (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI): Professional manual*. Odessa, FL: Psychological Assessment Resources.

Cromer TD, Hilsenroth MJ (2010). Patient personality and outcome in short-term psychodynamic psychotherapy. *J Nerv Ment Dis*, 198, 59-66.

Cuijpers P, Huibers M, Ebert DD, Koole SL, Andersson G (2013). How much psychotherapy is needed to treat depression? A metaregression analysis. *J Affect Disord*, 149, 1-13.

D'Zurilla TJ, Goldfried MR (1971). Problem solving and behavior modification. *J Abnorm Psychol*, 78, 107-126.

D'Zurilla TJ, Nezu AM (1999). *Problem-solving therapy: A social competence approach to clinical intervention*. New York, NY: Springer.

Davanloo H (1978). *Basic principles and techniques in short-term dynamic psychotherapy*. New York, NY: SP Medical and Scientific Books.

Davanloo H (1980). *Short-term dynamic psychotherapy*. New York, NY: Jason Aronson Inc.

Davanloo H (1986). Intensive short-term psychodynamic psychotherapy with highly resistant patietns. I. Handling resistance. *Int J Short-term Psychother*, 1, 107-133.

de Carufel FL, Piper WE (1988). Group psychotherapy or individual psychotherapy: patient characteristics as predictive factors. *Int J Group Psychother*, 38, 169-188.

de Jonghe F, de Maat S, Barber JP, Abbas A, Luyten P, Gomperts W, Swinkels J, Dekker J (2012). Designs for studying the effectiveness of long-term psychoanalytic treatments: balancing level of evidence and acceptability to patients. *J Am Psychoanal Assoc*, 60, 361-387.

Dekker JJ, Koelen JA, Van HL, Schoevers RA, Peen J, Hendriksen M, Kool S, van Aalst G, de Jonghe F (2008). Speed of action: the relative efficacy of short psychodynamic supportive psychotherapy and pharmacotherapy in the first 8 weeks of a treatment algorithm for depression. *J Affect Disord*, 109, 183-188.

de Maat S, de Jonghe F, de Kraker R, Leichsenring F, Abbass A, Luyten P, Barber JP, Rien V, Dekker J (2013). The current state of the empirical evidence for psychoanalysis: a meta-analytic approach. *Harv Rev Psychiatry*, 21, 107-137.

de Maat S, de Jonghe F, Schoevers R, Dekker J (2009). The effectiveness of long-term psychoanalytic therapy: a systematic review of empirical studies. *Harv Rev Psychiatry*, 17, 1-23.

Patient suitability for short-term and long-term psychotherapy

de Maat S, Dekker J, Schoevers R, van Aalst G, Gijsbers-van Wijk C, Hendriksen M, Kool S, Peen J, Van R, de Jonghe F (2008). Short psychodynamic supportive psychotherapy, antidepressants, and their combination in the treatment of major depression: a mega-analysis based on three randomized clinical trials. *Depress Anxiety*, 25, 565-574.

Derogatis LR (1983). *SCL-90-R Administration, Scoring & Procedures Manual-II*. Towson, MD: Clinical Psychometric Research.

Derogatis LR (2000). Symptom Checklist-90-Revised. In: A. J. Rush A., M. B. First, D. Blacker (Eds.). *Handbook of psychiatric measures*. Arlington, VA: American Psychiatric Association, 81-84.

Derogatis LR, Lipman RS, Covi L (1973). SCL-90: an outpatient psychiatric rating scale--preliminary report. *Psychopharmacol Bull*, 9, 13-28.

Derogatis LR, Rickels K, Rock AF (1976). The SCL-90 and the MMPI: a step in the validation of a new self-report scale. *Br J Psychiatry*, 128, 280-289.

DeRubeis RJ, Hollon SD, Amsterdam JD, Shelton RC, Young PR, Salomon RM, O'Reardon JP, Lovett ML, Gladis MM, Brown LL, Gallop R (2005). Cognitive therapy vs medications in the treatment of moderate to severe depression. *Arch Gen Psychiatry*, 62, 409-416.

de Shazer S (1991). *Putting difference to work*. New York, NY: W. W. Norton & Company.

de Shazer S, Berg IK, Lipchik E, Nunnally E, Molnar A, Gingerich W, Weiner-Davis M (1986). Brief therapy: Focused solution development. *Fam Process*, 25, 207-221.

Dewald PA (1964). *Psychotherapy: A dynamic approach*. New York, NY: Basic.

Doidge N (2001). Is psychoanalysis effective? *Econ Neurosci*, 3, 41-47.

Driessen E, Cuijpers P, de Maat SC, Abbass AA, de Jonghe F, Dekker JJ (2010). The efficacy of short-term psychodynamic psychotherapy for depression: a meta-analysis. *Clin Psychol Rev*, 30, 25-36.

Elkin I, Parloff MB, Hadley SW, Autry JH (1985). NIMH Treatment of Depression Collaborative Research Program. *Arch Gen Psychiatry*, 42, 305-316.

Elkin I, Shea MT, Watkins JT, Imber SD, Sotsky SM, Collins JF, Glass DR, Pilkonis PA, Leber WR, Docherty JP, Fiester SJ, Parloff MB (1989). National Institute of Mental Health Treatment of Depression Collaborative Research Program. General effectiveness of treatments. *Arch Gen Psychiatry*, 46, 971-982.

Emmelkamp PM (1980). Agoraphobics' interpersonal problems. Their role in the effects of exposure in vivo therapy. *Arch Gen Psychiatry*, 37, 1303-1306.

Eysenck HJ (1959). *The manual of the Maudsley Personality Inventory*. London: University of London Press.

Ezriel H (1952). Notes on psychoanalytic group therapy. II. Interpretation and research. *Psychiatry*, 15, 119-126.

Fenichel O (1954). Brief psychotherapy. In: H. Fenichel, D. Rapaport (Eds.). *The collected papers of Otto Fenichel*. New York, NY: W. W. Norton.

Fennell MJ, Teasdale JD (1987). Cognitive therapy for depression: individual differences and the process of change. *Cognitive Ther Res*, 11, 253-271.

Ferenczi S (1919). Technical difficulties in the analysis of a case of hysteria. In: J. Rickman (Ed.). *Further contributions to the theory and technique of psycho-analysis*. London: Hogarth Press, 1950, 189-197.

Ferenczi S (1920). The further development of an active therapy in psycho-analysis. In: J. Rickman (Ed.). *Further contributions to the theory and technique of psycho-analysis*. London: Hogarth Press, 1950, 198-217.

Ferenczi S (1925). Contra-indications to the 'active' psychoanalytic technique. In: J. Rickman (Ed.).

Further contributions to the theory and technique of psycho-analysis. London: Hogarth Press, 1950, 217-230.

Ferenczi S, Rank O (1925). *The development of psychoanalysis*. New York, NY: Nervous and Mental Disease Publishing Company.

Fisher D, Beutler LE, Williams OB (1999). Making assessment relevant to treatment planning: the STS Clinician Rating Form. Systemic Treatment Selection. *J Clin Psychol*, 55, 825-842.

Fleiss JL (1981). *Statistical methods for rates and proportions*. New York, NY: John Wiley & Sons.

Fonagy P (2002). *An open door review of outcome studies in psychoanalysis*. London: International Psychoanalytical Association.

Fournier JC, DeRubeis RJ, Shelton RC, Hollon SD, Amsterdam JD, Gallop R (2009). Prediction of response to medication and cognitive therapy in the treatment of moderate to severe depression. *J Consult Clin Psychol*, 77, 775-787.

Freud S (1895). Katharina. In: J. Strachey (Ed.). *The standard edition of the complete psychological works of Sigmund Freud*. London: Hogarth Press, 1955, 2, 125-134.

Freud S (1905a). On psychotherapy. In: J. Strachey (Ed.). *The standard edition of the complete psychological works of Sigmund Freud*. London: Hogarth Press, 1953, 7, 257-268.

Freud S (1905b). Fragment of an analysis of a case of hysteria. In: J. Strachey (Ed.). *The standard edition of the complete psychological works of Sigmund Freud*. London: Hogarth Press, 1953, 7, 7-122.

Freud S (1909). Notes upon a case of obsessional neurosis. In: J. Strachey (Ed.). *The standard edition of the complete psychological works of Sigmund Freud*. London: Hogarth Press, 1955, 10, 155-318.

Freud S (1919). Lines of advance in psycho-analytic therapy. In: J. Strachey (Ed.). *The standard edition of the complete psychological works of Sigmund Freud*. London: Hogarth Press, 1955, 17, 159-168.

Freud S (1923). The ego and the id In: J. Strachey (Ed.). *The standard edition of the complete psychological works of Sigmund Freud*. London: Hogarth Press, 1961, 19, 3-66.

Friedman RS, Lister P (1987). The current status of psychodynamic formulation. *Psychiatry*, 50, 126-141.

Fuerst RA (1938). Problems of short time psychotherapy. *Amer J Orthopsychiat*, 8, 260.

Gabbard GO (2004). *Long-term psychodynamic psychotherapy. A basic text*. Washington, DC: American Psychiatric Publishing.

Galatzer-Levy R, Bachrach H, Skolnikoff A, Waldron S (2000). *Does psychoanalysis work?* New Haven, CT: Yale University Press.

Garfield SL, Affleck DC (1961). Therapists' judgements concerning patients considered for psychotherapy. *J Consult Psychol*, 25, 505-509.

Gay ML, Hollandsworth JG, Galassi JP (1975). An assertiveness inventory. *J Counsel Psychol*, 22, 340-344.

Gillman RD (1965). Brief psychotherapy: a psychoanalytic view. *Am J Psychiatry*, 122, 601-611.

Gingerich WJ, Eisengart S (2000). Solution-focused brief therapy: a review of the outcome research. *Fam Process*, 39, 477-498.

GRADE Working Group, Atkins D, Best D, Briss PA, Eccles M, Falck-Ytter Y, Flottorp S, Guyatt GH, Harbour RT, Haugh MC, Henry D, Hill S, Jaeschke R, Leng G, Liberati A, Magrini N, Mason J, Middleton P, Mrukowicz J, O'Connell D, Oxman AD, Phillips B, Schunemann HJ, Edejer T, Varonen H, Vist GE, Williams JW, Jr., Zaza S (2004). Grading quality of evidence and strength of recommendations. *BMJ*, 328, 1490.

Grande T, Reiner D, Jakobsen T, Keller W, Krawietz B, Langer M, Oberbracht C, Stehle S, Stennes M, Rudolf G (2006). Differential effects of two forms of psychoanalytic therapy: Results of the

Heidelberg-Berlin study. *Psychother Res*, 16, 470-485.

Greenstein DK, Franklin ME, McGuffin P (1999). Measuring motivation to change: an examination of the University of Rhode Island Change Assessment Questionnaire (URICA) in an adolescent sample. *Psychother Theory Res Pract Train*, 36, 47-55.

Griffin DW, Bartholomew K (1994). The metaphysics of measurement: The case of adult attachment. In: K. Bartholomew, D. Perlman (Eds.). *Advances in personal relationships: Attachment processes in adulthood*. London: Jessica Kingsley, 5, 17-52.

Gwet KL (2012). *Handbook of inter-rater reliability*. Gaithersburg, MD: Advanced Analytics, LLC.

Haaga DA, DeRubeis RJ, Stewart BL, Beck AT (1991). Relationship of intelligence with cognitive therapy outcome. *Behav Res Ther*, 29, 277-281.

Hamilton M (1959). The assessment of anxiety states by rating. *Br J Med Psychol*, 32, 50-55.

Hamilton M (1960). A rating scale for depression. *J Neurol Neurosur Ps*, 23, 56-61.

Hardy GE, Cahill J, Shapiro DA, Barkham M, Rees A, Macaskill N (2001). Client interpersonal and cognitive styles as predictors of response to time-limited cognitive therapy for depression. *J Consult Clin Psychol*, 69, 841-845.

Hayes SC, Strosahl KD, Wilson KG (1999). *Acceptance and commitment therapy: An experiential approach to behavior change*. New York, NY: Wilson Guilford Press.

Hazan C, Shaver P (1987). Romantic love conceptualized as an attachment process. *J Pers Soc Psychol*, 52, 511-524.

Heiberg A (1975). Indications for psychotherapy in a psychiatric clinic population. A survey. *Psychother Psychosom*, 26, 156-166.

Heiberg A (1976). Indications for psychotherapy in a psychiatric clinic population. Reliability and validity of evaluations. *Psychother Psychosom*, 27, 18-25.

Heinonen E, Lindfors O, Laaksonen MA, Knekt P (2012). Therapists' professional and personal characteristics as predictors of outcome in short- and long-term psychotherapy. *J Affect Disord*, 138, 301-312.

Henry WP, Strupp HH, Butler SF, Schacht TE, Binder JL (1993). Effects of training in time-limited dynamic psychotherapy: changes in therapist behaviour. *J Consult Clin Psychol*, 61, 434-440.

Hersoug AG, Sexton HC, Høglend P (2002). Contribution of defensive functioning to the quality of working alliance and psychotherapy outcome. *Am J Psychother*, 56, 539-554.

Hilliard RB, Henry WP, Strupp HH (2000). An interpersonal model of psychotherapy: linking patient and therapist developmental history, therapeutic process, and types of outcome. *J Consult Clin Psychol*, 68, 125-133.

Høglend P (1993). Suitability for brief dynamic psychotherapy: psychodynamic variables as predictors of outcome. *Acta Psychiatr Scand*, 88, 104-110.

Høglend P, Engelstad V, Sørbye O, Heyerdahl O, Amlo S (1994). The role of insight in exploratory psychodynamic psychotherapy. *Br J Med Psychol*, 67 (Pt 4), 305-317.

Høglend P, Sørbye O, Sørlie T, Fossum A, Engelstad V (1992a). Selection criteria for brief dynamic psychotherapy: reliability, factor structure and long-term predictive validity. *Psychother Psychosom*, 57, 67-74.

Høglend P, Sørlie T, Heyerdahl O, Sørbye O, Amlo S (1993). Brief dynamic psychotherapy: patient suitability, treatment length and outcome. *J Psychother Pract Res*, 2, 230-241.

Høglend P, Sørlie T, Sørbye O, Heyerdahl O, Amlo S (1992b). Long-term changes after brief dynamic psychotherapy: symptomatic versus dynamic assessments. *Acta Psychiatr Scand*, 86, 165-172.

Patient suitability for short-term and long-term psychotherapy

Holi MM, Sammallahti PR, Aalberg VA (1998). A Finnish validation study of the SCL-90. *Acta Psychiatr Scand*, 97, 42-46.

Horowitz LM, Alden LE, Wiggins JS, Pincus AL (2000). *Inventory of the Interpersonal Problems. Manual.* San Antonio, TX: The Psychological Corporation.

Horowitz LM, Inouye D, Siegelman EY (1979). On averaging judges' ratings to increase their correlation with an external criterion. *J Consult Clin Psychol*, 47, 453-458.

Horowitz LM, Rosenberg SE, Baer BA, Ureno G, Villasenor VS (1988). Inventory of interpersonal problems: psychometric properties and clinical applications. *J Consult Clin Psychol*, 56, 885-892.

Horowitz MJ (1979). *States of mind*. New York, NY: Plenum Press.

Horowitz MJ, Marmar C, Weiss DS, DeWitt KN, Rosenbaum R (1984). Brief psychotherapy of bereavement reactions. The relationship of process to outcome. *Arch Gen Psychiatry*, 41, 438-448.

Huber D, Zimmermann J, Henrich G, Klug G (2012). Comparison of cognitive-behaviour therapy with psychoanalytic and psychodynamic therapy for depressed patients - a three-year follow-up study. *Z Psychosom Med Psychother*, 58, 299-316.

Huber D, Henrich G, Clarkin J, Klug G (2013). Psychoanalytic versus psychodynamic therapy for depression: a three-year follow-up study. *Psychiatry*, 76, 132-149.

Husby R (1985a). Short-term dynamic psychotherapy. III. A 5-year follow-up of 36 neurotic patients. *Psychother Psychosom*, 43, 17-22.

Husby R (1985b). Short-term dynamic psychotherapy. IV. Comparison of recorded changes in 33 neurotic patients 2 and 5 years after end of treatment. *Psychother Psychosom*, 43, 23-27.

Husby R, Dahl AA, Dahl CI, Heiberg AN, Olafsen OM, Weisaeth L (1985). Short-term dynamic psychotherapy. II. Prognostic value of characteristics of patients studied by a 2-year follow-up of 39 neurotic patients. *Psychother Psychosom*, 43, 8-16.

Hämäläinen J, Isometsä E, Sihvo S, Pirkola S, Kiviruusu O (2008). Use of health services for major depressive and anxiety disorders in Finland. *Depress Anxiety*, 25, 27-37.

Härkänen T, Knekt P, Virtala E, Lindfors O (2005). A case study in comparing therapies involving informative drop-out, non-ignorable non-compliance and repeated measurements. *Stat Med*, 24, 3773-3787.

Härkänen T, Arjas E, Laaksonen MA, Lindfors O, Haukka J, Knekt P (2013). Estimating efficacy in the presence of non-ignorable non-trial interventions in the Helsinki Psychotherapy Study. *Stat Methods Med Res*, Feb 1 [Epub ahead of print].

Imber SD, Pilkonis PA, Sotsky SM, Elkin I, Watkins JT, Collins JF, Shea MT, Leber WR, Glass DR (1990). Mode-specific effects among three treatments for depression. *J Consult Clin Psychol*, 58, 352-359.

Jarrett RB, Eaves GG, Grannemann BD, Rush AJ (1991a). Clinical, cognitive, and demographic predictors of response to cognitive therapy for depression: a preliminary report. *Psychiatry Res*, 37, 245-260.

Jarrett RB, Giles DE, Gullion CM, Rush AJ (1991b). Does learned resourcefulness predict response to cognitive therapy in depressed outpatients? *J Affect Disord*, 23, 223-229.

Jarrett RB, Kraft D, Doyle J, Foster BM, Eaves GG, Silver PC (2001). Preventing recurrent depression using cognitive therapy with and without continuation phase: a randomized clinical trial. *Arch Gen Psychiatry*, 58, 381-388.

Johnson LD, Miller SD (1994). Modification of depression risk factors: a solution-focused approach. *Psychother Theory Res Pract Train*, 31, 244-253.

Jørgensen CR, Hougaard E, Rosenbaum B, Valbak K, Rehfeld E (2000). The Dynamic Assessment Interview (DAI), interpersonal process measured by structural analysis of social behavior (SASB) and therapeutic outcome. *Psychother Res*, 10, 181-195.

Joutsenniemi K, Laaksonen MA, Knekt P, Haaramo P, Lindfors O (2012). Prediction of the outcome of short- and long-term psychotherapy based on socio-demographic factors. *J Affect Disord*, 141, 331-342.

Joyce PR, McKenzie JM, Carter JD, Rae AM, Luty SE, Frampton CM, Mulder RT (2007). Temperament, character and personality disorders as predictors of response to interpersonal psychotherapy and cognitive-behavioural therapy for depression. *Br J Psychiatry*, 190, 503-508.

Kampman M, Keijsers GPJ, Hoogduin CAL, Hendriks G-J (2008). Outcome prediction of cognitive behaviour therapy for panic disorder: initial symptom severity is predictive for treatment outcome, comorbid anxiety or depressive disorder, cluster C personality disorders and initial motivation are not. *Behav Cogn Psychoth*, 36, 99-112.

Keijsers G, Hoogduin C, Schaap C (1991). Motivatie voor psychotherapie: De ontwikkeling van een predictie-instrument [Motivation for psychotherapy: The development of a prediction instrument]. *Gedragstherapie*, 24, 195-208.

Keijsers GP, Hoogduin CA, Schaap CP (1994b). Predictors of treatment outcome in the behavioural treatment of obsessive-compulsive disorder. *Br J Psychiatry*, 165, 781-786.

Keijsers GPJ, Hoogduin CAL, Schaap CPDR (1994a). Prognostic factors in the behavioral treatment of panic disorder with and without agoraphobia. *Behav Ther*, 25, 689-708.

Keijsers GPJ, Schaap CPDR, Hoogduin CAL, Hoogsteyns B, de Kemp ECM (1999). Preliminary results of a new instrument to assess patient motivation for treatment in cognitive behavior therapy. *Behav Cogn Psychoth*, 27, 165-179.

Kernberg OF (1981). Structural interviewing. *Psychiatr Clin North Am*, 4, 169-195.

Kernberg OF (1996). A psychoanalytic theory of personality disorders. In: J. F. Clarkin, M. F. Lenzenweger (Eds.). *Major theories of personality disorder*. New York, NY: The Guilford Press, 106-140.

Kernberg OF, Burstein ED, Coyne L, Appelbaum A, Horwitz L, Voth H (1972). Psychotherapy and psychoanalysis: the final report of the Menninger Foundation's Psychotherapy Research Project. *Bull Menninger Clin*, 36, 1-275.

Klerman GL, Weissman MM, Rounsaville BJ (1984). *Interpersonal psychotherapy of depression*. New York, NY: Basic Books.

Klerman GL, Rounsaville BJ, Chevron E, Neu C, Weissman MM (1979). *Manual for short-term interpersonal therapy (IPT) of depression* (fourth revision). New Haven-Boston Collaborative Depression Project. Unpublished manuscript.

Knekt P, Laaksonen M, Lindfors O (2010). *Helsingin Psykoterapiatutkimus - psykoterapioiden vaikuttavuus viiden vuoden seurannassa*. Helsinki: Terveyden ja hyvinvoinnin laitos (THL), Raportti 33.

Knekt P, Lindfors O (2004). *A randomized trial of the effect of four forms of psychotherapy on depressive and anxiety disorders. Design, methods, and results on the effectiveness of short-term psychodynamic psychotherapy and solution-focused therapy during a one-year follow-up*. Helsinki: The Social Incurance Institution, Studies in social security and health 77.

Knekt P, Lindfors O, Härkänen T, Välikoski M, Virtala E, Laaksonen MA, Marttunen M, Kaipainen M, Renlund C (2008a). Randomized trial on the effectiveness of long-and short-term psychodynamic psychotherapy and solution-focused therapy on psychiatric symptoms during a 3-year follow-up. *Psychol Med*, 38, 689-703.

Knekt P, Lindfors O, Laaksonen MA, Raitasalo R, Haaramo P, Järvikoski A (2008b). Effectiveness of short-term and long-term psychotherapy on work ability and functional capacity--a randomized clinical trial on depressive and anxiety disorders. *J Affect Disord*, 107, 95-106.

Knekt P, Lindfors O, Laaksonen MA, Renlund C, Haaramo P, Härkänen T, Virtala E (2011a). Quasi-experimental study on the effectiveness of psychoanalysis, long-term and short-term psychotherapy on psychiatric symptoms, work ability and functional capacity during a 5-year follow-up. *J Affect Disord*, 132, 37-47.

Knekt P, Lindfors O, Renlund C, Sares-Jäske L, Laaksonen MA, Virtala E (2011b). Use of auxiliary psychiatric treatment during a 5-year follow-up among patients receiving short- or long-term psychotherapy. *J Affect Disord*, 135, 221-230.

Knekt P, Lindfors O, Sares-Jäske L, Virtala E, Härkänen T (2013). Randomized trial on the effectiveness of long- and short-term psychotherapy on psychiatric symptoms and working ability during a 5-year follow-up. *Nord J Psychiatry*, 67, 59-68.

Knight RP (1937). Application of psychoanalytic concepts in psychotherapy. *Bull Menninger Clin*, 1, 99.

Konsensuskokous (2006). *Psykoterapia.* Suomalainen lääkäriseura Duodecim ja Suomen Akatemia. Vammala: Vammalan kirjapaino.

Korchin SJ, Schuldberg D (1981). The future of clinical assessment. *Am Psychol*, 36, 1147-1158.

Kordy H, von Rad M, Senf W (1988). Time and its relevance for a successful psychotherapy. *Psychother Psychosom*, 49, 212-222.

Krippendorf K (1970). Bivariate agreement coefficient for reliability of data. In: F. Borgatta, G. W. Bohrnstedt (Eds.). *Sociological methodology*. San Francisco, CA: Jossey-Bass, 139-150.

Kronström K, Salminen JK, Hietala J, Kajander J, Vahlberg T, Markkula J, Rasi-Hakala H, Karlsson H (2009). Does defense style or psychological mindedness predict treatment response in major depression? *Depress Anxiety*, 26, 689-695.

Kronström K, Salminen JK, Hietala J, Kajander J, Vahlberg T, Markkula J, Rasi-Hakala H, Karlsson H (2011). Personality traits and recovery from major depressive disorder. *Nord J Psychiatry*, 65, 52-57.

Laaksonen M (2010). *Population Attributable Fraction (PAF) in epidemiologic follow-up studies*. Helsinki: Terveyden ja hyvinvoinnin laitos (THL), Tutkimus 34.

Laaksonen MA, Sirkiä C, Knekt P, Lindfors O (2014). Self-reported immature defense style as a predictor of outcome in short- and long-term psychotherapy. *Brain and Behavior*, 4, 495-503.

Lake B (1985). Concept of ego strength in psychotherapy. *Br J Psychiatry*, 147, 471-478.

Lambert MJ, Okiishi JC, Finch AE, Johnson LD (1998). Outcome assessment: from conceptualization to implementation. *Prof Psychol Res Pr*, 29, 63–70.

Lazarus RS, Folkman S (1984). *Stress, appraisal, and coping*. New York, NY: Springer.

Lee J (1981). Covariance adjustment of rates based on the multiple logistic regression model. *J Chronic Dis*, 34, 415-426.

Leichsenring F, Abbass A, Luyten P, Hilsenroth M, Rabung S (2013). The emerging evidence for long-term psychodynamic therapy. *Psychodyn Psychiatry*, 41, 361-384.

Leichsenring F, Leibing E (2007). Psychodynamic psychotherapy: a systematic review of techniques, indications and empirical evidence. *Psychol Psychother*, 80, 217-228.

Leichsenring F, Rabung S (2008). Effectiveness of long-term psychodynamic psychotherapy: a meta-analysis. *JAMA*, 300, 1551-1565.

Leichsenring F, Rabung S (2011). Long-term psychodynamic psychotherapy in complex mental disorders: update of a meta-analysis. *Br J Psychiatry*, 199, 15-22.

Patient suitability for short-term and long-term psychotherapy

Lewis CC, Simons AD, Kim HK (2012). The role of early symptom trajectories and pretreatment variables in predicting treatment response to cognitive behavioral therapy. *J Consult Clin Psychol*, 80, 525-534.

Linden A (2013). Assessing regression to the mean effects in health care initiatives. *BMC Med Res Methodol*, 13, 119.

Lindfors O, Knekt P, Heinonen E, Virtala E (2014). Self-concept and quality of object relations as predictors of outcome in short- and long-term psychotherapy. *J Affect Disord*, 152-154, 202-211.

Lindfors O, Knekt P, Virtala E, Laaksonen MA (2012). The effectiveness of solution-focused therapy and short- and long-term psychodynamic psychotherapy on self-concept during a 3-year follow-up. *J Nerv Ment Dis*, 200, 946-953.

Lindsley O, Skinner BF, Solomon HC (1953). *Studies in behavior therapy (Status Report I)*. Walthama, MA: Metropolitan State Hospital.

Luborsky L, Singer B, Luborsky L (1975). Comparative studies of psychotherapies. Is it true that "everywon has one and all must have prizes"? *Arch Gen Psychiatry*, 32, 995-1008.

Luty SE, Carter JD, McKenzie JM, Rae AM, Frampton CM, Mulder RT, Joyce PR (2007). Randomised controlled trial of interpersonal psychotherapy and cognitive-behavioural therapy for depression. *Br J Psychiatry*, 190, 496-502.

Mahrer AR (1983). *Experiential psychotherapy: Basic practices*. Ottawa: University of Ottawa Press.

Mahrer AR (1989). *How to do experiential psychotherapy: a manual for practitioners*. Ottawa: University of Ottawa Press.

Mahrer AR (1996). *The complete guide to experiential psychotherapy*. New York, NY: John Wiley & Sons.

Malan D (1986). Beyond interpretation: Initial evaluation and technique in Short-Term Dynamic Psychotherapy. Part II. *Int J Short-term Psychother*, 1, 83-106.

Malan DH (1963). *A study of brief psychotherapy*. London: Tavistock Press.

Malan DH (1976a). *The frontier of brief psychotherapy: An example of the convergence of research and clinical practice*. New York, NY: Plenum Medical Book Company.

Malan DH (1976b). *Toward the validation of dynamic psychotherapy. A replication*. New York, NY: Plenum Medical Book Company.

Mann J (1973). *Time-limited psychotherapy*. Cambridge, MA: Harvard University Press.

Marshall MB, Zuroff DC, McBride C, Bagby RM (2008). Self-criticism predicts differential response to treatment for major depression. *J Clin Psychol*, 64, 231-244.

Maruish ME (1994). *The use of psychological testing for treatment planning and outcome assessment*. Hillsdale, NJ: Lawrence Erlbaum & Associates, Inc

McBride C, Atkinson L, Quilty LC, Bagby RM (2006). Attachment as moderator of treatment outcome in major depression: a randomized control trial of interpersonal psychotherapy versus cognitive behavior therapy. *J Consult Clin Psychol*, 74, 1041-1054.

McBride C, Zuroff DC, Ravitz P, Koestner R, Moskowitz DS, Quilty L, Bagby RM (2010). Autonomous and controlled motivation and interpersonal therapy for depression: moderating role of recurrent depression. *Br J Clin Psychol*, 49, 529-545.

McCallum M, Piper WE (1997). The Psychological Mindedness Assessment Procedure. In: M. McCallum, W. E. Piper (Eds.). *Psychological mindedness: A contemporary understanding*. Mahwah, NJ: Erlbaum.

McConnaughy EA, DiClemente CC, Prochaska JO, Velicer WF (1989). Stages of change in psychotherapy: A follow-up report. *Psychother Theory Res Pract*, 26, 494-503.

McConnaughy EA, Prochaska JO, Velicer WF (1983). Stages of change in psychotherapy:

Measurement and sample profiles. *Psychother Theory Res Pract*, 20, 368-375.

McCullough JP (2000). *Treatment for chronic depression: Cognitive Behavioral Analysis System of Psychotherapy (CBASP)*. New York, NY: Guilford Publications.

McGuire MT (1965). The process of short-term insight psychotherapy. *J Nerv Ment Dis*, 141, 83-94.

Menninger K (1958). *Theory of psychoanalytic technique*. New York, NY: Basic Books.

Migon HdS, Gamerman D (1999). *Statistical inference: an integrated approach*. London: Arnold.

Monsen JT, Monsen K, Solbakken OA, Hansen RS (2008). *The Affect Consciousness Interview (ACI) and the Affect Consciousness Scales (ACS): Instructions for the interview and rating*. Oslo: University of Oslo.

Murphy GE, Simons AD, Wetzel RD, Lustman PJ (1984). Cognitive therapy and pharmacotherapy, singly and together, in the treatment of depression. *Arch Gen Psychiatry*, 135, 525-533.

Myhr G, Talbot J, Annable L, Pinard G (2007). Suitability for short-term cognitive-behavioral therapy. *J Cogn Psychother*, 21, 334-345.

Nielsen G, Havik OE, Barth K, Haver B, Mølstad E, Rogge H, Skatun M (1984). Evaluering av pasienter for dynamisk korttids psykoterapi: En project illustrasjon [Evaluation of patients for dynamic short-term psychotherapy: A project illustration]. *Nord Psychol*, 36, 65-74.

Nieuwsma JA, Trivedi RB, McDuffie J, Kronish I, Benjamin D, Williams JW (2012). Brief psychotherapy for depression: a systematic review and meta-analysis. *Int J Psychiatry Med*, 43, 129-151.

OPD Task Force (2001). *Operationalized psychodynamic diagnostics: foundations and manual*. Seattle, WA: Hogrefe & Huber.

Paitich D, Crawford G (1970). *A multiple-choice version of the WAIS Vocabulary*. Toronto: Clarke Institute of Psychiatry. Unpublished manuscript.

Paivio SC, Bahr LM (1998). Interpersonal problems, working alliance, and outcome in short-term experiential therapy. *Psychother Res*, 8, 392-407.

PDM Task Force (2006). *Profile for mental functioning. Psychodynamic diagnostic manual*. Silver Spring, MD: Alliance of Psychoanalytic Organizations.

Perry JC (1990). *Defense Mechanism Rating Scales (DMRS)*. Cambridge, MA: Author.

Perry JC, Høglend P (1998). Convergent and discriminant validity of overall defensive functioning. *J Nerv Ment Dis*, 186, 529-535.

Perry S, Cooper AM, Michels R (1987). The psychodynamic formulation: its purpose, structure, and clinical application. *Am J Psychiatry*, 144, 543-550.

Persons JB, Burns DD, Perloff JM (1988). Predictors of dropout and outcome in cognitive therapy for depression in a private practice setting. *Cognitive Ther Res*, 12, 557-575.

Persson G, Alström JE (1983). A scale for rating suitability for insight-oriented psychotherapy. *Acta Psychiatr Scand*, 68, 117-125.

Peterson C, Semmel A, von Baeyer C, Abramson LY, Metalsky GI, Seligman MEP (1982). The Attributional Style Questionnaire. *Cognitive Ther Res*, 6, 287-300.

Piper WE, de Carufel FL, Szkrumelak N (1985). Patient predictors of process and outcome in short-term individual psychotherapy. *J Nerv Ment Dis*, 173, 726-733.

Piper WE, Joyce AS, McCallum M, Azim HF (1998). Interpretive and supportive forms of psychotherapy and patient personality variables. *J Consult Clin Psychol*, 66, 558-567.

Pirkola SP, Isometsä E, Suvisaari J, Aro H, Joukamaa M, Poikolainen K, Koskinen S, Aromaa A, Lönnqvist JK (2005). DSM-IV mood-, anxiety- and alcohol use disorders and their comorbidity in the Finnish general population--results from the Health 2000 Study. *Soc Psych Psych Epid*, 40, 1-10.

Pumpian-Mindlin E (1953). Considerations in the selection of patients for short-term therapy. *Am J Psychother*, 7, 641-652.

Puschner B, Kraft S, Bauer S (2004). Interpersonal problems and outcome in outpatient psychotherapy: findings from a long-term longitudinal study in Germany. *J Pers Assess*, 83, 223-234.

Rachman S (1997). The evolution of cognitive behaviour therapy. In: D. Clark, C. G. Fairburn, M. G. Gelder (Eds.). *Science and practice of cognitive behaviour therapy.* Oxford: Oxford University Press, 1–26.

Renner F, Jarrett RB, Vittengl JR, Barrett MS, Clark LA, Thase ME (2012). Interpersonal problems as predictors of therapeutic alliance and symptom improvement in cognitive therapy for depression. *J Affect Disord*, 138, 458-467.

Resick PA, Nishith P, Weaver TL, Astin MC, Feuer CA (2002). A comparison of cognitive-processing therapy with prolonged exposure and a waiting condition for the treatment of chronic posttraumatic stress disorder in female rape victims. *J Consult Clin Psychol*, 70, 867-879.

Ripley HS, Wolf S, Wolff HG (1948). Treatment in a psychosomatic clinic; preliminary report. *J Am Med Assoc*, 138, 949-951.

Rizvi SL, Vogt DS, Resick PA (2009). Cognitive and affective predictors of treatment outcome in Cognitive Processing Therapy and Prolonged Exposure for posttraumatic stress disorder. *Behav Res Ther*, 47, 737-743.

Rogers CR (1951). *Client-centered therapy: Its current practice, implications, and theory.* Boston, MA: Houghton Mifflin.

Rogers CR (1957). The necessary and sufficient conditions of therapeutic personality change. *J Consult Psychol*, 21, 95-103.

Rogers CR (1961). *On becoming a person: A therapist's view of psychotherapy.* Boston, MA: Houghton Mifflin.

Rosenbaum B, Selzer MA, Valbak K, Hougaard E, Sommerlund B (1997). The Dynamic Assessment Interview: testing the psychodynamic assessment variables. *Acta Psychiatr Scand*, 95, 531-538.

Rosenbaum M (1980). A schedule for assessing self control behaviors: Preliminary findings. *Behav Ther*, 11, 109-121.

Rosenbaum M, Friedlander J, Kaplan SM (1956). Evaluation of the results of psychotherapy. *Psychosom Med*, 18, 113-132.

Rosenberg M (1965). *Society and the adolescent child*. Princeton, NJ: Princeton University Press.

Roshanaei-Moghaddam B, Pauly MC, Atkins DC, Baldwin SA, Stein MB, Roy-Byrne P (2011). Relative effects of CBT and pharmacotherapy in depression versus anxiety: is medication somewhat better for depression, and CBT somewhat better for anxiety? *Depress Anxiety*, 28, 560-567.

Rothman KJ, Greenland S, Lash TL (2008). *Modern epidemiology* (3rd ed.). Philadelphia, PA: Lippincott Williams & Wilkins.

Rubin DB (1987). *Multiple imputation for nonresponse in surveys*. New York, NY: John Wiley & Sons.

Saatsi S, Hardy GE, Cahill J (2007). Predictors of outcome and completion status in cognitive therapy for depression. *Psychother Res*, 17, 185-195.

Safran JD, Segal ZV, Vallis TM, Shaw BF, Samstag LW (1993). Assessing patient suitability for short-term cognitive therapy with an interpersonal focus. *Cognitive Ther Res*, 17, 23-28.

Sandell R, Blomberg J, Lazar A, Carlsson J, Broberg J, Schubert J (2000). Varieties of long-term outcome among patients in psychoanalysis and long-term psychotherapy. A review of findings in the Stockholm Outcome of Psychoanalysis and Psychotherapy Project (STOPP). *Int J Psychoanal*, 81 (Pt 5), 921-942.

SAS Institute Inc. (2008). *SAS/STAT User's Guide, Version 9.2*. Cary, NC: SAS Institute Inc.

Schutz WC (1973). Marital attitudes evaluation. In: J. W. Pfeiffer, R. Heslin (Eds.). *Instrumentation in human relations training*. Iowa City, IA: Iowa University Association.

Searle SR (1971). *Linear models*. New York, NY: Wiley.

Seeman M (1967). Powerlessness and knowledge: A comparative study of alienation and learning. *Sociometry*, 30, 105-123.

Seligman ME (1995). The effectiveness of psychotherapy. The Consumer Reports study. *Am Psychol*, 50, 965-974.

Seligman ME, Abramson LY, Semmel A, von Baeyer C (1979). Depressive attributional style. *J Abnorm Psychol*, 88, 242-247.

Selzer MA, Koenigsberg HW, Kernberg OF (1987). The initial contract in the treatment of borderline patients. *Am J Psychiatry*, 144, 927-930.

Shedler J (2010). The efficacy of psychodynamic psychotherapy. *Am Psychol*, 65, 98-109.

Sherer M, Maddux J, Mercandante B, Prentice-Dunn S, Jacobs B, Rogers R (1982). The Self-Efficacy Scale: construction and validation. *Psychol Rep*, 51, 663-671.

Shill MA, Lumley MA (2002). The Psychological Mindedness Scale: factor structure, convergent validity and gender in a non-psychiatric sample. *Psychol Psychother*, 75, 131-150.

Shipley WC (1940). A self-administering scale for measuring intellectual impairment and deterioration. *J Psychol*, 9, 371-377.

Sifneos PE (1965). Seven years' experience with short-term psychotherapy. In: *Selected lectures, 6th International Congress of Psychotherapy, London, 1964*. New York, NY: S. Krager.

Sifneos PE (1966). Psychoanalytically oriented short-term dynamic or anxiety-provoking psychotherapy for mild obsessional neuroses. *Psychiatr Q*, 40, 271-282.

Sifneos PE (1967). Two different kinds of psychotherapy of short duration. *Am J Psychiatry*, 123, 1069-1074.

Sifneos PE (1968a). "The motivational process"--a selection and prognostic criterion for psychotherapy of short duration. *Psychiatr Q*, 42, 271-279.

Sifneos PE (1968b). Learning to solve emotional problems: a controlled study of short-term anxiety-provoking psychotherapy. *Int Psychiatry Clin*, 6, 87-99.

Sifneos PE (1972). *Short-term psychotherapy and emotional crisis*. Cambridge, MA: Harvard University Press.

Sifneos PE (1978). Short-term anxiety provoking psychotherapy. In: H. Davanloo (Ed.). *Short-term dynamic psychotherapy*. New York, NY: Spectrum, 35-42.

Sifneos PE (1979). *Short-term dynamic psychotherapy: Evaluation and technique*. New York, NY: Plenum Press.

Sigal JJ, Paris J, Kardos M, Zimmerman G, Buonvino M (1999). Evaluation of some criteria used to select patients for brief psychodynamic therapy. *Psychother Psychosom*, 68, 193-198.

Sihvo S, Hämäläinen J, Kiviruusu O, Pirkola S, Isometsä E (2006). Treatment of anxiety disorders in the Finnish general population. *J Affect Disord*, 96, 31-38.

Sillanpää M, Andlin-Sobocki P, Lönnqvist J (2008). Costs of brain disorders in Finland. *Acta Neurol Scand*, 117, 167-172.

Simons AD, Lustman PJ, Wetzel RD, Murphy GE (1985). Predicting response to cognitive therapy of depression: The role of learned resourcefulness. *Cognitive Ther Res*, 9, 79-89.

Simons AD, Gordon JS, Monroe SM, Thase ME (1995). Toward and integration of psychologic, social, and biologic factors in depression: effects on outcome and course of cognitive therapy. *J Consult Clin Psychol*, 63, 369-377.

Small L (1971). *The brief psychotherapies*. New York, NY: Brunner/Mazel.

Smit Y, Huibers MJ, Ioannidis JP, van Dyck R, van Tilburg W, Arntz A (2012). The effectiveness of long-term psychoanalytic psychotherapy--a

meta-analysis of randomized controlled trials. *Clin Psychol Rev*, 32, 81-92.

Socarides CW (1954). On the usefulness of extremely brief psychoanalytic contacts. *Psychoanal Rev*, 41, 340-346.

Solbakken OA, Abbass A (2014). Implementation of an intensive short-term dynamic treatment program for patients with treatment-resistant disorders in residential care. *BMC Psychiatry*, 14, 12.

Solbakken OA, Hansen RS, Havik OE, Monsen JT (2012). Affect integration as a predictor of change: affect consciousness and treatment response in open-ended psychotherapy. *Psychother Res*, 22, 656-672.

Sotsky SM, Glass DR, Shea MT, Pilkonis PA, Collins JF, Elkin I, Watkins JT, Imber SD, Leber WR, Moyer J, Oliveri ME (1991). Patient predictors of response to psychotherapy and pharmacotherapy: findings in the NIMH Treatment of Depression Collaborative Research Program. *Am J Psychiatry*, 148, 997-1008.

Spangler DL, Simons AD, Monroe SM, Thase ME (1997). Response to cognitive-behavioral therapy in depression: effects of pretreatment cognitive dysfunction and life stress. *J Consult Clin Psychol*, 65, 568-575.

Spielberger CD, Sydeman SJ (1994). State-trait anxiety inventory and state-trait anger expression inventory. In: C. D. Spielberger, S. J. Sydeman, E. M. Maruish (Eds.). *The use of psychological testing for treatment planning and outcome assessment*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Steel Z, Marnane C, Iranpour C, Chey T, Jackson JW, Patel V, Silove D (2014). The global prevalence of common mental disorders: a systematic review and meta-analysis 1980-2013. *Int J Epidemiol*, 43, 476-493.

Steketee G, Siev J, Fama JM, Keshaviah A, Chosak A, Wilhelm S (2011). Predictors of treatment outcome in modular cognitive therapy for obsessive-compulsive disorder. *Depress Anxiety*, 28, 333-341.

Stoler N (1963). Client likeability: a variable in the study of psychotherapy. *J Consult Psychol*, 27, 175-178.

Strupp HH, Binder JL (1984). *Psychotherapy in a new key*. New York, NY: Basic Books.

Stuart S, Robertson M (2003). *Interpersonal psychotherapy: A clinician's guide*. London: Edward Arnold.

Sullivan PL, Miller C, Smelser W (1958). Factors in length of stay and progress in psychotherapy. *J Cons Psychol*, 22, 1-9.

Terman LM (1916). *The Measurement of intelligence: An explanation of and a complete guide to the use of the stanford revision and extension of the Binet-Simon intelligence scale*. Boston, MA: Houghton Mifflin.

Trijsburg RW, van t' Spijker A, Van HL, Hesselink AJ, Duivenvoorden HJ (2000). Measuring overall defensive functioning with the Defense Style Questionnaire: a comparison of different scoring methods. *J Nerv Ment Dis*, 188, 432-439.

Valbak K (2004). Suitability for psychoanalytic psychotherapy: a review. *Acta Psychiatr Scand*, 109, 164-178.

Valbak K, Rosenbaum B, Hougaard E (2004). Suitability for psychoanalytic psychotherapy: validation of the Dynamic Assessment Interview (DAI). *Acta Psychiatr Scand*, 109, 179-186.

Van HL, Dekker J, Peen J, Abraham RE, Schoevers R (2009). Predictive value of self-reported and observer-rated defense style in depression treatment. *Am J Psychother*, 63, 25-39.

Van HL, Hendriksen M, Schoevers RA, Peen J, Abraham RA, Dekker J (2008). Predictive value of object relations for therapeutic alliance and outcome in psychotherapy for depression: an exploratory study. *J Nerv Ment Dis*, 196, 655-662.

Vaslamatzis G, Markidis M, Katsouyanni K (1989). Study of the patients' difficulties in ending brief

psychoanalytic psychotherapy. *Psychother Psychosom*, 52, 173-178.

Vaslamatzis G, Verveniotis S (1985). Early dropouts in brief dynamic psychotherapy. *Psychother Psychosom*, 44, 205-210.

Verbeke G, Molenberghs G (1997). *Linear mixed models in practice: an SAS-oriented approach*. New York, NY: Springer.

Wampold BE (2001). *The great psychotherapy debate: models, methods, and findings*. Mahwah, NJ: Erlbaum.

Wampold BE, Budge SL, Laska KM, Del Re AC, Baardseth TP, Fluckiger C, Minami T, Kivlighan DM 2nd, Gunn W (2011). Evidence-based treatments for depression and anxiety versus treatment-as-usual: a meta-analysis of direct comparisons. *Clin Psychol Rev*, 31, 1304-1312.

Wampold BE, Minami T, Baskin TW, Callen Tierney S (2002). A meta-(re)analysis of the effects of cognitive therapy versus 'other therapies' for depression. *J Affect Disord*, 68, 159-165.

Wechsler D (1955). *Wechsler Adult Intelligence Scale Manual*. New York, NY: Psychological Corporation.

Weissman AN (1979). *The Dysfunctional Attitudes Scale: A validation study*. Philadelphia, PA: University of Pennsylvania.

Weissman MM, Bothwell S (1976). Assessment of social adjustment by patient self-report. *Arch Gen Psychiatry*, 33, 1111-1115.

WHO (2000). Cross-national comparisons of the prevalences and correlates of mental disorders. WHO International Consortium in Psychiatric Epidemiology. *Bull World Health Organ*, 78, 413-426.

Williams GC, Freedman ZR, Deci EL (1998). Supporting autonomy to motivate patients with diabetes for glucose control. *Diabetes Care*, 21, 1644-1651.

Wolberg LR (1965). *Short-term psychotherapy*. New York, NY: Grune and Stratton.

Wolberg LR (1977). *The technique of psychotherapy*. New York, NY: Grune & Stratton.

Wolitzky-Taylor KB, Arch JJ, Rosenfield D, Craske MG (2012). Moderators and non-specific predictors of treatment outcome for anxiety disorders: a comparison of cognitive behavioral therapy to acceptance and commitment therapy. *J Consult Clin Psychol*, 80, 786-799.

Yates AJ (1970). *Behavior Therapy*. New York, NY: Wiley.

Zeidner M, Endler NSE (1996). *Handbook of coping - theory, research, applications*. New York, NY: Wiley.

Zuckerman DM, Prusoff BA, Weissman MM, Padian NS (1980). Personality as a predictor of psychotherapy and pharmacotherapy outcome for depressed outpatients. *J Consult Clin Psychol*, 48, 730-735.

Zuroff DC, Koestner R, Moskowitz DS, McBride C, Marshall M, Bagby RM (2007). Autonomous motivation for therapy: A new common factor in brief treatments for depression. *Psychother Res*, 17, 137-147.

9 SUPPLEMENTARY TABLES

Supplementary table 1A. Associations between psychological suitability measures within *nature of problems domain* and psychiatric symptoms.[1]

Author[2]	Data[2]	Predictor		Outcome		Therapy	Follow-up	N	Strength of association			Method	Adjustment	Result[1]
		Measure	Type	Measure	Type				Min	Med	Max			
P1: FOCUS														
Short-term therapies														
Brodaty et al. 1982	3	1	1	V3: 1	3	PP(I)	0	18		NR^0		1	-	0/1
							12	18		NR^0				0/1
Husby et al. 1985	5	2, 3	1	V3: 2	3	PP(I)	24	36	-0.31^0	-0.07^0	0.18^0	1	-	0/2
Husby 1985a							60	33	-0.05^0	0.04^0	0.12^0			0/2
Høglend 1993	11	4	2	V3: 4	3	PP(I)	24	40		0.08^0		1	-	0/1
							48	43		0.16^0				0/1
Sigal et al. 1999	19	5	1	V3: 5, 6	2	PP(I)	0	43	NR^0	NR^0	NR^0	1	-	0/2
Myhr et al. 2007	27	6	1	V3: 7	6	CBT	0	113		0.15^0		1	-	0/1

[1] See Appendix 2 for detailed description of terminology and definitions.
[2] See Table 6 for detailed description of the datasets and individual studies.

Supplementary table 1B. Associations between psychological suitability measures within *ego strength domain* and psychiatric symptoms.[1]

Author[2]	Data[2]	Predictor Measure	Predictor Type	Outcome Measure	Outcome Type	Therapy	Follow-up	N	Min	Med	Max	Method	Adjustment	Result
P2: MODULATION OF AFFECTS														
Short-term therapies														
Brodaty et al. 1982	3	1	1	V3: 1	3	PP(I)	0	18		NR^0		1	-	0/1
Sigal et al. 1999	19	2, 3	1, 2	V3: 5, 6	2	PP(I)	12	18		0.63^{**}		1	-	1/1
							0	43	NR^0	NR^0	NR^0			0/4
Myhr et al. 2007	27	4	1	V3: 7	6	CBT	0	113		0.26^{**}		1	-	1/1
Cromer & Hilsenroth 2010	33	5-7	1	V3: 3, 8	6	PP(I)	0	71	NR^0	NR^0	NR^0	2	CDPS	0/6
		8	2						0.07^0	0.11^0	0.14^0	1	-	0/2
Long-term therapies														
Solbakken et al. 2012	41	9	2	V3: 1	3	PP(I)	0	153		NA^*		2	AGE, GENDER, TIME, ALL, PD	1/1
P3: FLEXIBILITY OF INTERACTION														
Short-term therapies														
Husby et al. 1985	5	1	1	V3: 2	3	PP(I)	24	36		0.61^{**}		1	-	1/1
Husby 1985a							60	33		0.42^*				1/1
Høglend 1993	11	2	2	V3: 4	3	PP(I)	24	40		0.17^0		1	-	0/1
							48	43		0.09^0				0/1
Myhr et al. 2007	27	3	1	V3: 7	6	CBT	0	113		0.19^*		1	-	1/1
Cromer & Hilsenroth 2010	33	4	1	V3: 3, 8	6	PP(I)	0	71	NR^0	NR^0	NR^0	2	CDPS	0/2
P4: SELF-CONCEPT IN RELATION TO EGO IDEAL														
Short-term therapies														
Horowitz et al. 1984	4	1	2	V3: 1	3	PP(I)	0	51		0.17^0		1	BAS	0/1

[1] See Appendix 2 for detailed description of terminology and definitions.
[2] See Table 6 for detailed description of the datasets and individual studies.

Supplementary table 1C. Associations between psychological suitability measures within *self-observing capacity domain* and psychiatric symptoms.[1]

Author[2]	Data[2]	Predictor Measure	Predictor Type	Outcome Measure	Outcome Type	Therapy	Follow-up	N	Strength of association Min	Med	Max	Method	Adjustment	Result
P5: RESPONSE TO TRIAL INTERPRETATION														
Short-term therapies														
Brodaty et al. 1982	3	1	1	V3: 1	3	PP(I)	0	18	-	NR[0]	-	1	-	0/1
							12	18	-	NR[0]	-			0/1
P6: REFLECTIVE ABILITY														
Short-term therapies														
Brodaty et al. 1982	18	1	1	V3: 1	3	PP(I)	0	18		NR[0]		1	-	0/1
							12	18		0.45*				1/1
Piper et al. 1998	18	2	2	V1: 4, V2: 11, V3: 1	2	PP(I)	0	64	0.11[0]	0.15[0]	0.26*	1	-	1/3
						PP(S)		66	0.14[0]	0.21[0]	0.23[0]			0/3
Myhr et al. 2007	27	8	1	V3: 7	6	CBT	0	113		0.18[0]		1	-	0/1
Kronström et al. 2009	32	3	2	V1: 3	2	PP(I)	0	19		NR[0]		1	-	0/1
					3					NR[0]		2		0/1
					2					NR[0]		1		0/1
					3		8	19		NR[0]		2		0/1
Cromer & Hilsenroth 2010	33	4	1	V3: 3, 8	6	PP(I)	0	71	0.26*	0.27*	0.27*	2	-	2/2
		5, 6	1						NR[0]	NR[0]	NR[0]	2	CDPS	0/4
		7	2						0.17[0]	0.21	0.24*		-	1/2
P7: MOTIVATION														
Short-term therapies														
Brodaty et al. 1982	3	1	1	V3: 1	3	PP(I)	0	18		NR[0]		1	-	0/1
							12	18		NR[0]				0/1
Horowitz et al. 1984	4	2	2	V3: 1	3	PP(I)	0	48	NR[0]	-0.14[0]	-0.14[0]	1	BAS	0/1
Husby et al. 1985 Husby 1985a	5	3	1	V3: 2	3	PP(I)	24	36		0.05[0]		1	-	0/1
							60	33		0.34*				1/1
Sotsky et al. 1991	10	10	2	V1: 3	2	CBT	0	59[3]		NR[0]		2	-	0/1
								37[4]		NR[0]				0/1
						IPT	0	61[3]		NR[0]		2	-	0/1
								47[4]		NR[0]				0/1

Study				Dataset		Treatment	n	N				k	Predictors[1]	Ratio
Høglend 1993	11	6	2	V1: 3+4	1	CBT	0	59[3]			NR[0]	2	-	0/1
								37[4]			NR[0]			0/1
						IPT	0	61[3]			NR[0]	2	-	0/1
								47[4]			NR[0]			0/1
Keijsers et al. 1994a	13	7	2	V3: 4	3	PP(I)	24	40			0.07[0]	1	-	0/1
							48	43			0.15[0]			0/1
				V2: 7d	2	BT	2	60	0.20[*]		0.21[*]	1	-	1/1
				V2: 7a, b, 8	4			53-		0.20[*]	0.24[**]			1/1
				V2: 7c	1			60					CAQ, Rlp	3/3
				V2: 7d	1			53		NR[*]		2	CAQ, Rlp, HDRS	1/1
				V2: 9	1			60		NR[*]			CAQ, HDRS, PP	1/1
								60		NR[*]				1/1
Keijsers et al. 1994b	14	7	2	V2: 9, 10	4	BT	0	40	0.10[0]	0.17	0.24[*]	1	Rlp, Rlt, HDRS	1/2
				V2: 9	1			40		NR[0]			Rlt, HDRS, DUR	0/1
				V2: 10	1			40		NR[0]		2	-	0/1
Sigal et al. 1999	19	11-13	1, 2	V3: 5, 6	2	PP(I)	0	43	NR[0]	NR[0]	NR[0]	1	-	0/6
Zuroff et al. 2007	24	8, 9	2	V1: 3	1	CBT	0	36	NA[0]	NA	NA[*]	2	GRP	1/2
						IPT	0	30	NA[0]	NA	NA[*]			1/2
				V1: 3, 4	2	CBT	0	36	NA[0]	NA	NA[*]	2	BAS, GRP, ALL, EACH OTHER	2/4
						IPT	0	30	NA[0]	NA	NA[*]			2/4
Myhr et al. 2007	27	14-16	1	V3: 7	6	CBT	0	113	0.17[0]	0.19[0]	0.21[*]	1	-	1/3
Kampman et al. 2008	28	7	2	V2: 7c, 10, 12, 13a, b	2	CBT	0	114-	NA[0]	NA[0]	NA[0]	2	BAS	0/5
				V2: 7d	1			129						
McBride et al. 2010	34	8, 9	2	V1: 4	2	IPT	0	129	NR[0]	NA[*]	NR[*]	2	BAS, MS, AGE, DR, EACH OTHER	0/1
							0	71		NR			BAS	1/2
Steketee et al. 2011	35	4	2	V2: 15	3	CT	0	28	NA[*]	NA[*]		2	BAS	1/1
							12	28	NA[0]	NA[0]				0/1
Lewis et al. 2012	36	5A-D	2	V1: 4	2	CBT	0	173	NA[0]	NA[0]	NA[*]	2	BAS, BAI, BDI1-5, GAF, COM, BHS, DAS, AGE, SS, TT, GENDER, MEDS, EACH OTHER	1/4

[1] See Appendix 2 for detailed description of terminology and definitions.
[2] See Table 6 for detailed description of the datasets and individual studies.
[3] Intention to treat sample: patients randomized to treatments.
[4] Completer sample: patients who completed the treatment.

Supplementary table 1D. Associations between psychological suitability measures within *intelligence domain* and psychiatric symptoms.[1]

Author[2]	Data[2]	Predictor		Outcome		Therapy	Follow-up	N	Strength of association			Method	Adjustment	Result
		Measure	Type	Measure	Type				Min	Med	Max			
P8: INTELLIGENCE														
Short-term therapies														
Husby et al. 1985	5	1	1	V3: 2	3	PP(I)	24	36		0.34^{*}		1	-	1/1
Husby 1985a							60	33		0.44^{**}				1/1
Høglend et al. 1994	11	2	1	V3: 4	3	PP(I)	24	40		NR^{0}		1	-	0/1
							48	43		NR^{0}				0/1
Haaga et al. 1991	8	3, 4	2	V1: 4	4	CT	0	76	0.07^{0}	0.12^{0}	0.15^{0}	1	-	0/2
				V2: 5				30	0.17^{0}	0.21^{0}	0.24^{0}			0/2
Fournier et al. 2009	30	5	2	V1: 3	2	CT	0	60		NA^{*}		2	BAS	1/1
Rizvi et al. 2009	31	6	2	V2: 14	2	BT	0	56		NA^{0}		2	-	0/1
						CBT	0	56		NA^{0}				0/1

[1] See Appendix 2 for detailed description of terminology and definitions.
[2] See Table 6 for detailed description of the datasets and individual studies.

Supplementary table 1E. Associations between psychological suitability measures within *intrapsychic and interpersonal behavior domain* and psychiatric symptoms.[1]

Author[2]	Data[2]	Predictor Measure	Type	Outcome Measure	Type	Therapy	Follow-up	N	Min	Med	Max	Method	Adjustment	Result[1]
P9: COPING STYLES AND COGNITIVE SKILLS													
Short-term therapies													
Zuckerman et al. 1980	2	14	2	V1: 3, 4	2	IPT	0	NR	NA[0]	NA[0]	NA[0]	2	BAS	0/2
							12	NR	NA[0]	NA[0]	NA[0]			0/2
Simons et al. 1985	6	13	2	V1: 4	2	CT	0	19		0.53[*]	NR[0]	2	BAS	1/1
		1, 10-12							NR[0]	NR[0]			BAS, SCS	0/4
Jarrett et al. 1991a	9	1, 16	2	V1: 3, 4	2	CT	0	37	0.01[NR]	0.19[NR]	0.36[NR]	1	-	NR/4
		1, 16							NR[0]	0.24[0]	0.32[*]	2	BAS	1/4
		1							0.22[0]	0.25[0]	0.28[0]		BAS, MS	0/2
Jarrett et al. 199b	9	13	2	V1: 3, 4	2	CT	0	37	NA[0]	NA[0]	NA[0]	2	BAS	0/2
Sotsky et al. 1991	10	9	2	V1: 3	2	CBT	0	59[3]		NR[NR]		2	BAS, DUR, AD, ONSET, HIST	NR/1
						CBT	0	37[4]		NR[NR]			BAS, DUR, AD, ONSET, HIST, MS	NR/1
						IPT	0	61[3]		NR[NR]		2	BAS, SD, MS, IS, SSR	NR/1
						IPT	0	47[4]		NR[NR]				NR/1
						CBT	0	59[3]		NR[NR]				NR/1
				V1: 3+4	1	CBT	0	37[4]		NR[NR]		2	BAS, DUR, AD, ONSET, HIST	NR/1
						IPT	0	61[3]		NR[NR]		2	BAS, SD, MS, IS, SSR	NR/1
						IPT	0	47[4]		NR[NR]				NR/1
						CBT	0	NR						
Blatt et al. 1995	10	2, 3	2	V1: 3, 4	2	CBT	0	NR	NA[0]	NA	NA[*]	2	BAS, MS	4/8
				V3: 1, 4		IPT	0	NR	NA[0]	NA	NA[*]			4/8
Simons et al. 1995	15	1	2	V1: 3, 4	2	CBT	0	53	NR[0]	NR	NR[*]	2	BAS	1/2
				V1: 3	1	IPT				NR[0]				0/1
Spangler et al. 1997	16	4, 5, 17, 18	2	V1: 3, 4	2	CBT	0	53	NA[0]	NA[0]	NA[*]	2	BAS, LE, DAS, ASQ	2/8
Hardy et al. 2001	21	6-8	2	V1: 4	2	CT	0	24	0.14[0]	0.33[0]	0.39[0]	1	-	0/3
									0.01[0]	0.04[0]	0.18[0]	2	BAS	0/3
Clark et al. 2003b	23	1, 16, 17, 20	2	V1: 7	2	CT	0	100	0.01[0]	0.07[0]	0.10[0]	1	-	0/4

177

Patient suitability for short-term and long-term psychotherapy

Supplementary table 1E. Associations between psychological suitability measures within *intrapsychic and interpersonal behavior domain* and psychiatric symptoms.[1]

Author[2]	Data[2]	Predictor Measure	Predictor Type	Outcome Measure	Type	Therapy	Follow-up	N	Min	Med	Max	Method	Adjustment	Result[1]
Fournier et al. 2009	30	2, 3, 12, 15, 21	2	V1:3	2	CBT	0	60	NA^0	NA^0	NA^0	2	-	0/5
Carter et al. 2011	26	1	2	V1:8	5	CBT+IPT	0	177		NR*		2	CHILD, LOGIC, SUCCESS, GRP SUCCESS*GRP	0/1
Lewis et al. 2012	36	1, 12	2	V1:4	2	CBT	0	173	NA^0	NA^0	NA^0	2	BAS, BAI, GAF, SS, TT, MEDS, COM, SS, AGE GENDER, SOCS, EACH OTHER	0/2

P10. DEFENSE STYLES

Short-term therapies

Author[2]	Data[2]	Predictor Measure	Predictor Type	Outcome Measure	Type	Therapy	Follow-up	N	Min	Med	Max	Method	Adjustment	Result[1]
Hersoug et al. 2002	22	1, 2	2	V3:1, 4	2	PP(I)	0	39	NA^0	NA^0	NA^0	2	-	0/4
Myhr et al. 2007	27	16	1	V3:7	6	CBT	0	113	NR^0	0.25**	NR***	1	-	1/1
Van et al. 2009	29	6-15	2	V1:3	1	PP(S)	0	81	NR^0	NR^0	NR***	2	BAS, MED	3/10
		2-5						69	NR^0	NR	NR			2/4
		2, 3, 5						81	NR^0	0.28*	0.44*	2	-	2/3
		6						69	NR*	NR*	NR*			1/1
		7, 8						69	NR*	NR^0	NR*			2/2
Kronström et al. 2009	32	3-5	2	V1:3	2	PP(I)	0	19	NR^0	NR^0	NR^0	1	GENDER EACH OTHER	0/3
					3			19	NR^0	NR^0	NR^0	2		0/3
					2		8	19	NR^0	NR^0	NR^0	1	-	0/3
					3			19	NR^0	NR^0	NR^0	2		0/3

P11. PERSONALITY TRAITS

Short-term therapies

Author[2]	Data[2]	Predictor Measure	Predictor Type	Outcome Measure	Type	Therapy	Follow-up	N	Min	Med	Max	Method	Adjustment	Result[1]
Zuckerman et al. 1980	2	28, 29, 37	2	V1:3, 4	2	IPT	0	NR	NA^0	NA^0	NA^0	2	BAS	0/6
							12	NR	NA^0	NA^0	NA^0	2		0/6
Brodaty et al. 1982	3	1	1	V3:1	3	PP(I)	0	18	NR^0	NR^0	NR^0	1	-	0/1
							12	18	NR^0	NR^0	NR^0			0/1
Sigal et al. 1999	19	2	2	V3:5, 6	2	PP(I)	0	43	NR^0	NR^0	NR^0	1	-	0/2
Clark et al. 2003a	23	3-17	2	V1:7	2	CT	0	108	NR^0	NR^0	0.16^0	1	-	0/15
Clark et al. 2003b	23	3-8	2	V1:7	2	CT	0	100	0.00^0	0.15^0	0.17^0	1	-	0/6

Study	ID	Ref		V		Therapy		N					Predictor	Ratio
Marshall et al. 2008	24	25, 26	2	V1: 3	2	CBT	0	37	0.04[0]	0.18[0]	0.32[0]	1	-	0/2
									NR[0]	NR[0]	NR[0]	2	BAS	0/2
									0.11[0]	0.22[0]	0.33[0]	1	-	0/2
Blom et al. 2007b	25	31-35	2	V1: 3	1	IPT	0	35	NR[0]	NR	NR[*]	2	BAS	1/2
									NA[0]	NA[0]	NA[0]	2	BAS, DUR, MEDC	0/5
Joyce et al. 2007	26	18-24	1	V1: 8	2	CBT	0	34	NA[0]	NA[0]	NA[0]	1	-	0/5
		18-24			5	IPT	0	34	0.07[0]	0.17[0]	0.22[*]			1/7
		19, 20					0	80	-0.02[0]	0.22[0]	0.37[***]	2	EACH OTHER	3/7
		19					0	87	0.21	0.28[*]	0.35[*]		AVO, SCH	2/2
										0.25[*]				
Rizvi et al. 2009	31	36	2	V2: 14	2	BT	0	56		NA[0]	NA[0]	2	-	1/1
						CBT	0	56		NA[0]	NA[0]	2	-	0/1
						PP(I)	0	19	0.00[0]	0.14[0]	0.22[0]	2	-	0/1
Kronström et al. 2011	32	18-24	1	V1: 3	2	PP(I)	0	71	NR[0]	NR[0]	NR[0]	2	-	0/7
Cromer & Hilsenroth 2010	33	27	1	V3: 3, 8	6	PP(I)	0					2	CDPS	0/2
Wolitzky-Taylor et al. 2012	38	30	2	V2: 16	2	ACT	0	33		NR[*]		2	BAS	1/1
						CBT	0	41		NR[*]		2	BAS	1/1
						ACT	6	20		NR[*]		2	BAS	1/1
						CBT	6	29		NR[*]		2	BAS	1/1
						ACT	12	16		NR[*]		2	BAS	1/1
						CBT	12	21		NR[*]		2	BAS	1/1

P12: INTERPERSONAL RELATIONSHIPS

Short-term therapies

Study	ID	Ref		V		Therapy		N					Predictor	Ratio
Emmelkamp 1980	1	1, 2	2	V1: 1; V2: 1a-f, 2a-f, 3a, b, 4	2	BT	0	17	NA[0]	NA[0]	NA[*]	2	-	1/32
Husby et al. 1985, Husby 1985a	5	3	1	V3: 2	2	PP(I)	1	17	NA[0]	NA[0]	NA[0]	1	-	0/32
							24	36		0.30[0]				0/1
							60	33		0.19[0]				0/1
Sotsky et al. 1991	10	10	2	V1: 3	2	CBT	0	59[3]		NR[0]		2	-	0/1
								37[4]		NR[0]				0/1
						IPT	0	61[3]		NR[NR]		2	BAS	NR/1
								47[4]		NR[0]				0/1
				V1: 3+4	1	CBT	0	59[3]		NR[0]			-	0/1
								37[4]		NR[0]				
						IPT		61[3]	0.35[NR]	0.35[NR]		2	BAS	NR/1
								47[4]		0.35[NR]				NR/1

Supplementary table 1E. Associations between psychological suitability measures within *intrapsychic and interpersonal behavior domain* and psychiatric symptoms.[1]

Author[2]	Data[2]	Predictor Measure	Predictor Type	Outcome Measure	Outcome Type	Therapy	Follow-up	N	Min	Med	Max	Method	Adjustment	Result
Høglend et al. 1992a	11	4	1	V3: 4	3	PP(I)	24	34		NR^*		2	AGE, SC	1/1
Høglend 1993	11	4	1	V3: 4	3	PP(I)	24	40		0.35^*		1	-	1/1
								40		0.30^*		2	BAS, PD	0/1
							48	43		0.17^0		1	-	0/1
								43		0.10^0		2	BAS, PD	0/1
Høglend et al. 1993	11	4	1	V3: 4	3	PP(I)	24	40		NR^*		2	-	1/1
Paivio & Bahr 1998	17	16–25	2	V3: 1	4	ET	48	43		NR^0		1		0/1
							0	33	NR^0	NR^0	NR^0			0/10
Piper et al. 1998	18	5	2	V1: 4 / V2: 11 / V3: 1	2	PP(I)	0	64	0.14^0	0.26^*	0.38^{**}	1		2/3
						PP(S)	0	66	-0.07^0	0.00^0	0.02^0			0/3
Sigal et al. 1999	19	6, 7	1, 2	V3: 5, 6	2	PP(I)	0	43	NR^0	NR^0	NR^0	1	-	0/4
Hilliard et al. 2000	20	8	2	V3: 1	4	PP(I)	0	50		0.27^*		1	-	1/1
Hardy et al. 2001	21	29, 30	2	V1: 4	2	CT	0	24	0.57^{**}	0.59^{**}	0.60^{**}	1	BAS	2/2
									0.51^{**}	0.57^{**}	0.62^{**}		BAS	2/2
		30								0.56^{**}		2		1/1
Saatsi et al. 2007	21	31, 32, 33 / 31	2	V1: 4	2	CT	0	97	NR^0	NR^*	0.24^*	2	BAS	1/3
Clark et al. 2003b	23	9, 15	2	V1: 7	2	CT	0	97		0.10^0			BAS, ALL	0/1
McBride et al. 2006	24	34	2	V1: 3, 4	2	CBT	0	100	0.12^0	0.15^0	0.18^0	1	-	0/2
								37^3	NR^0	NR^0	NR^0	2	BAS, AAVO	0/2
								28^4	NR^0	NR^0	NR^0			0/2
								37^3	NR^0	NR^0	NR^0		BAS, AAVO, OCPD, APD	0/2
								28^4	NR^0	NR^0	NR^0		BAS, AAVO	0/2
						IPT		36^3	NR^0	NR^0	NR^0	2	BAS, AAVO, OCPD, APD	0/2
								27^4	NR^0	NR^0	NR^0		BAS, AAVO	0/2
								36^3	NR^0	NR^0	NR^0		BAS, ANX	
								27^4	NR^0	NR^0	NR^0			
		35		V1: 3, 4	2	CBT	0	37^3	NR^*	NR^*	NR^*	2	BAS, AANX, OCPD, APD	2/2
								28^4	NR^*	NR^*	NR^*			2/2
								37^3	NR^*	NR^*	NR^*			2/2
								28^4	NR^*	NR^*	NR^*			2/2

Study						Treatment		N				n	Comparison	Ratio
Carter et al. 2011	26	11-14	2	V1:7	5	IPT	0	36[3]	NR[0]	NR[0]	NR[0]	2	BAS, ANX	0/2
		11						27[4]	NR[0]	NR	NR[*]		BAS, AANX, OCPD, APD	1/2
								36[3]	NR[0]	NR[0]	NR[0]		-	0/2
								27[4]	NR[0]	NR[0]	NR[0]			0/2
Myhr et al. 2007	27	36	1	V3:7	6	CBT+IPT	0	177	NR[0]	NR[0]	NR[0]	2	-	0/4
						CBT+IPT	0	177		0.14[0]		1	GRP	0/1
Van et al. 2008	29	37-46	2	V1:3	6	CBT	0	113	NR[0]	0.23[*]	NR[*]	2	-	1/1
		36, 44, 45, 46				PP(S)	0	81	NR[0]	NR[0]		2	-	4/10
								81	NA[0]	NA[0]	0.26[~]		-	1/4
Cromer & Hilsenroth 2010	33	48	2	V3: 3, 8	6	PP(I)	0	71	0.11[0]	0.14[0]	0.17[0]	1	-	0/2
		47	1					71	NR[0]	NR[0]	NR[0]	2	CDPS	0/2
Renner et al. 2012	37	26-28	2	V1:3	2	CT	0	523	NA[0]	NA[0]	NA[*]	2	-	1/3
Long-term therapies														
Jørgensen et al. 2000	39	49, 50	2	V3: 1	3	PP(I)	0	16	0.54[*]	0.56[*]	0.57[*]	1	-	2/2
Puschner et al. 2004	40	24, 25	2	V3: 9	6	CBT	0	NR	0.00[0]	0.02[0]	0.03[0]	2	EACH OTHER	0/2
						PP(I)	0	NR	0.03[0]	0.06	0.09[*]			1/2

[1] See Appendix 2 for detailed description of terminology and definitions.
[2] See Table 6 for detailed description of the datasets and individual studies.
[3] Intention to treat sample: patients randomized to treatments.
[4] Completer sample: patients who completed the treatment.

Supplementary table 1F. Associations between psychological suitability measures within *overall suitability domain* and psychiatric symptoms.[1]

Author[2]	Data[2]	Predictor Measure	Type	Outcome Measure	Type	Therapy	Follow-up	N	Strength of association Min	Med	Max	Method	Adjustment	Result
P13: OVERALL SUITABILITY														
Short-term therapies														
Brodaty et al. 1982	3	1	2	V3: 1	3	PP(I)	0 12	18 18		NR[0] 0.55*		1	-	0/1 1/1
Alpher et al. 1990	7	2	2	V3: 1, 3	4	PP(I)	0	25	0.45*	0.51*	0.56**	2	-	2/2
Safran et al. 1993	12	3	2	V1: 4-6 V2: 6 V3: 1	2	CBT	0	32-42	0.29[0]	0.39*	0.46**	1	BAS	4/5
Myhr et al. 2007	27	3	2	V3: 7	6	CBT	0	113 111		0.30*** 0.20*		1 2	- MS, ANX	1/1 1/1
Cromer & Hilsenroth 2010	33	2	2	V3: 3, 8	6	PP(I)	0	71	0.20[0]	0.22[0]	0.23[0]	1	-	0/2
Long-term therapies														
Jørgensen et al. 2000	39	4	2	V3: 1	3	PP(I)	0	16		0.18[0]		1	-	0/1

[1] See Appendix 2 for detailed description of terminology and definitions.
[2] See Table 6 for detailed description of the datasets and individual studies.

10 APPENDICES

APPENDIX 1. CLASSIFICATION OF THE SUITABILITY ITEMS ON THE PUBLISHED PSYCHOTHERAPY SUITABILITY ASSESSMENT SCALES

SUITABILITY ITEMS ON THE SCALE[1]	CLASSIFICATION[2]

1A-C. Heiberg 1975; Husby et al. 1985, Husby 1985a, Barth et al. 1988a, b; Høglend et al. 1992a

1. Focus[3]　　　　　　　　　　　　　　　　P1.　Focus
   a. "Is the chief complaint well circumscribed?"
   b. "Is the chief problem well circumscribed?"

2. "Does the patient interact emotionally with evaluator during interview?"[4]　　　　P3.　Flexibility of interaction
   a. Flexible emotional interaction in interview
   b. Flexible intellectual interaction in interview

3. Motivation[5]　　　　　　　　　　　　　　P7.　Motivation
   a. Ability to recognize that symptoms are psychological
   b. Talent for introspection and honesty in self-report
   c. Willingness to participate actively in treatment
   d. Active curiosity about oneself
   e. Desire to change
   f. Realistic expectations of treatment results
   g. Willingness to make reasonable sacrifices

4. "Does the patient have above average intelligence / problem-solving capacity?"　　P8.　Intelligence

5. "Is there a history of at least one meaningful relationship with another person?"　　P12. Interpersonal relationships

6. Severity of symptoms and complaints[6]　　Other

Suitability score: sum of 5/6 item values　　P13. Overall suitability

2. Brodaty et al. 1982

1. Focus　　　　　　　　　　　　　　　　　P1.　Focus
2. Ability to emote　　　　　　　　　　　　P2.　Modulation of affects
3. Ability to accept interpretations　　　　　P5.　Trial interpretation
4. Psychological mindedness　　　　　　　　P6.　Reflective ability
5. Motivation　　　　　　　　　　　　　　　P7.　Motivation
6. Therapist optimism (patient attractiveness and likability)　P11. Personality traits
Total predictor variable (TPV): sum of 6 item values　　P13. Overall suitability

Patient suitability for short-term
and long-term psychotherapy

3. Persson & Alström 1983

1.	Possibility of psychodynamic formulation with a circumscribed focus	P1.	Focus
2.	Ability to tolerate anxiety	P2.	Modulation of affects
3.	Ability to experience positive emotions	P2.	Modulation of affects
4.	Ability to experience negative emotions	P2.	Modulation of affects
5.	Ability to describe emotional reactions	P2.	Modulation of affects
6.	Ability to express positive emotions	P2.	Modulation of affects
7.	Ability to express negative emotions	P2.	Modulation of affects
8.	Ability to form trusting relationship with the interviewer	P3.	Flexibility of interaction
9.	Evaluation of own symptoms as psychologically possible to understand or not	P6.	Reflective ability
10.	Ability to take a reasonable amount of responsibility	P7.	Motivation
11.	Willingness to get rid of the symptoms	P7.	Motivation
12.	Willingness to work actively with oneself	P7.	Motivation
13.	Intelligence	P8.	Intelligence
14.	Perseverance when confronted with troubles	P9.	Coping styles
15.	Ability to modify environment or change it	P9.	Coping styles
16.	Tendency to seek substitute satisfactions	P10.	Defense styles
17.	Self-confidence	P11.	Personality traits
18.	Demanding attitude	P11.	Personality traits
19.	Ability to form and maintain emotional ties	P12.	Interpersonal relationships
20.	Influence of environmental factors on the symptoms	Other	
21.	Variability of the symptoms during the last year	Other	
22.	Content of leisure time in relation to what is desired	Other	
23.	Secondary gain	Other	
24.	Attitude of significant others towards symptoms	Other	

4. Buckley et al. 1984

1.	Ego strength (5 items)	P2.	Modulation of affects
2.	Coping Styles (8 items)	P9.	Coping Styles
3.	Ego Defense Rating (16 items)	P10.	Defense styles
4.	Personality Trait and Inferred Conflict (16 items)	P11.	Personality Traits
5.	Object Relations (5 items)	P12.	Interpersonal relationships
6.	Adaptive functioning (6 items)	Other	
7.	Anxiety-Depression-Anger (3 items)	Other	

5. Piper et al. 1985

1.	Focality	P1.	Focus
2.	Psychodynamic formulation	P1.	Focus
3.	Affect appropriateness	P2.	Modulation of affects

4. Affect control	P2. Modulation of affects
5. Interaction	P3. Flexibility of interaction
6. Response to interpretation	P5. Trial interpretation
7. Psychological mindedness	P6. Reflective ability
8. Motivation to work in individual therapy	P7. Motivation
9. Motivation to work in group therapy	P7. Motivation
10. Defensive style	P10. Defense styles
11. Likability	P11. Personality traits
12. Object choice	P12. Interpersonal relationships
13. Object type	P12. Interpersonal relationships
14. Sexual relationships	P12. Interpersonal relationships
15. Psychosexual conflict	P12. Interpersonal relationships

6. Vaslamatzis & Verveniotis 1985, Vaslamatzis et al. 1989

1. The therapist has got a psychodynamic view of the patient's main problem	P1. Focus
2. Patient's problem is circumscribed	P1. Focus
3. Patient's reaction to the therapist pointing out transference feelings	P5. Trial interpretation
4. Motivation for change in the patient	P7. Motivation
5. Experience of a meaningful relationship with a significant other	P12. Interpersonal relationships
Suitability score: sum of 5 item values	P13. Overall suitability

7. Alpher et al. 1990

1. "integrates affect"	P2. Modulation of affects
2. "differentiates affects"	P2. Modulation of affects
3. "perceives affective aspects of problems"	P2. Modulation of affects
4. "collaborates therapeutically"	P3. Flexibility of interaction
5. "appears introspective"	P6. Reflective ability
6. "manifests insight"	P6. Reflective ability
7. "manifests verbal fluency"	P6. Reflective ability
8. "offers positive relationship"	P11. Personality traits
9. "differentiates interpersonal events"	P12. Interpersonal relationships
CDPS total score: average of 9 item values	P13. Overall suitability

8A-B. Safran et al. 1993; Myhr et al. 2007

1. Focality	P1. Focus
2. Awareness and differentiation of emotions	P2. Modulation of affects
3. Alliance potential: in-session evidence	P3. Flexibility of interaction
4. Accessibility of automatic thoughts	P6. Reflective ability
5. Acceptance of personal responsibility for change	P7. Motivation

6. Compatibility with cognitive rationale	P7. Motivation
7. General optimism/pessimism regarding therapy[7]	P7. Motivation
8. Security operations	P10. Defense styles
9. Alliance potential: out-of-session evidence	P12. Interpersonal relationships
10. Chronicity of problems	Other
Mean SSCT: average of 9/10 item values	P13. Overall suitability

9. Rosenbaum et al. 1997

1. Tolerance of frustration	P2. Modulation of affects
2. Ability to reveal and contain affect	P2. Modulation of affects
3. Ability to respond to confrontation	P5. Trial interpretation
4. Psychological mindedness	P6. Reflective ability
5. Capacity for self-observation	P6. Reflective ability
6. Motivation	P7. Motivation
7. Confidence in the treatment	P7. Motivation
8. Capacity for empathy	P11. Personality traits
9. Interviewer's attraction towards working with the patient	P11. Personality traits
10. Global suitability	P13. Overall suitability
Suitability score: sum of 10 item values	P13. Overall suitability

10. Fisher et al. 1999[8]

1. Coping style (patient externalization and internalization)	P9. Coping styles
2. Patient resistance potential	P11. Personality traits
3. Patient subjective distress	Other

11. Sigal et al. 1999[9]

1. Identifiable focus	P1. Focus
2. Anxiety verbalized and controlled	P2. Modulation of affects
3. Affect mobilized and recognized	P2. Modulation of affects
4. Recognizes feelings	P2. Modulation of affects
5. Maintains self without supervision	P4. Self concept
6. Can talk about fantasies	P6. Reflective ability
7. Realistic expectations for change	P7. Motivation
8. Searches for understanding	P7. Motivation
9. Thinks therapy won't help	P7. Motivation
10. Thinks will feel better after therapy	P7. Motivation
11. Takes no responsibility for complaints	P7. Motivation
12. No problem ending therapy	P7. Motivation
13. Will terminate prematurely	P7. Motivation
14. Termination will be easy	P7. Motivation
15. Transference manifestations	P12. Interpersonal relationships
16. Accepts paradox in interpersonal relations	P12. Interpersonal relationships

 Patient suitability for short-term
and long-term psychotherapy

17. Shows concern for other's need	P12. Interpersonal relationships
18. Talks to no one regularly	P12. Interpersonal relationships
19. Talks to at least one person regularly	P12. Interpersonal relationships
20. Loss experienced and talked about	Other
21. Mourned a former relationship	Other
22. No severe character pathology	Other
23. Borderline pathology	Other
24. Severe behavior disorder	Other
25. Difficulty attending school or job	Other
26. Attends school or works	Other
27. No previous therapy	Other
28. Successful previous therapy	Other
29. Tolerates absences	Other
30. At least one creative activity	Other
31. Has never worked	Other
32. Recent precipitant	Other

[1] See Table 3 for more detailed description of the scales.

[2] See Table 1.

[3] Two aspects are considered in the evaluation of focus. Husby et al. 1985, Husby1985a, Barth et al. 1988a, b, and Høglend et al. 1992a separated these two aspects into two separate items.

[4] Høglend et al. 1992a evaluated flexibility of both emotional and intellectual interaction.

[5] Seven aspects are considered in the evaluation of patient's motivation. Høglend et al. 1992a separated these seven aspects into seven separate items.

[6] Only in Høglend et al. 1992a.

[7] Only in Myhr et al. 2007.

[8] The 226 items of the Systematic Treatment Selection (STS) Clinician Rating Form were grouped into three factors and individuals items were not described.

[9] Only 32 of the 39 items of the Checklist for Brief Psychodynamic Therapy were reported.

APPENDIX 2. TERMINOLOGY

Terminology applied in Table 6 and Supplementary tables 1A-F in alphabetical order:

ADJUSTMENT
Factors for which the association between the predictor and the outcome was adjusted.

AANX = Attachment anxiety

AAVO = Attachment avoidance

AD = Affective disorder

AGE = Age

ALL = Alliance (patient-rated)

ANX = Anxiety disorder

APD = Avoidant personality disorder symptom count

ASQ = Attributional style (Attributional Style Questionnaire)

AVO = Avoidant personality disorder symptoms

BAI = Baseline anxiety symptoms (Beck Anxiety Inventory)

BAS = Baseline symptoms (of the symptom outcome measure)

BDI1-5 = Rate of early symptom change (Beck Depression Inventory, sessions 1-5)

CAQ = Catastrophic agoraphobic cognitions (Agoraphobic Cognitions Questionnaire)

CDPS = Capacity for Dynamic Process Scale (other 8 items of the scale)

CHILD = Childhood reasons for therapy (Reasons for Depression Questionnaire)

COM = Comorbidity

DAS = Dysfunctional attitudes (Dysfunctional Attitudes Scale)

DR = Depression recurrence

DUR = Duration of complaints / current episode of depression

EACH OTHER = Psychological predictors of interest adjusted for one another

GAF = Global functioning (Global Assessment of Functioning)

GENDER = Gender

GRP = Treatment group

HDRS = Baseline depressive symptoms (Hamilton Depression Rating Scale)

BHS = Hopelessness (Beck Hopelessness Scale)

HIST = History of previous episodes of depression

IS = Interpersonal sensitivity

LE = Life event

LOGIC = Logicalness of therapy (Treatment credibility)

MED = Addition of medication after two months of treatment (yes / no)

MEDC = Medical consumption prior to treatment (low / high)

MEDS = Medication status

MS = Marital status

NS = Number of sessions

OCPD = Obsessive-compulsive personality disorder symptom count

ONSET = Age at onset of psychiatric disorder
PD = Personality disorder (presence / absence)
PP = Personality psychopathology (Personality Disorders Questionnaire Revised)
RIp = Quality of therapeutic relationship (Relationship Inventory, patient-rated)
RIt = Quality of therapeutic relationship (Relationship Inventory, therapist-rated)
SC = Social class
SCH = Schizoid personality disorder symptoms
SCS = Learned resourcefulness (Self-Control Schedule)
SD = Social dysfunction
SOCS = Readiness to change (Stages of Change Scale)
SS = Student status
SSR = Satisfaction with social relationships
SUCCESS = Successfulness of therapy (Treatment credibility)
TIME = Follow-up time
TT = Therapist training

AGE
Mean age (range) if not otherwise mentioned.

AUTHOR & DATA

Short-term therapies
 1 = Emmelkamp 1980
 2 = Zuckerman et al. 1980
 3 = Brodaty 1982
 4 = Horowitz et al. 1984
 5 = Husby et al. 1985, Husby 1985a
 6 = Simons et al. 1985
 7 = Alpher et al. 1990
 8 = Haaga et al. 1991
 9 = Jarrett et al. 1991a, b
10 = Sotsky et al. 1991, Blatt et al. 1995
11 = Høglend et al. 1992a, 1993, 1994, Høglend 1993
12 = Safran et al. 1993
13 = Keijsers et al. 1994a
14 = Keijsers et al. 1994b
15 = Simons et al. 1995
16 = Spangler et al. 1997
17 = Paivio & Bahr 1998
18 = Piper et al. 1998
19 = Sigal et al. 1999
20 = Hilliard et al. 2000

21 = Hardy et al. 2001, Saatsi et al. 2007
22 = Hersoug et al. 2002
23 = Clark et al. 2003a, b
24 = McBride et al. 2006, Zuroff et al. 2007, Marshall et al. 2008
25 = Blom et al. 2007b
26 = Joyce et al. 2007, Carter et al. 2011
27 = Myhr et al. 2007
28 = Kampman et al. 2008
29 = Van et al. 2008, 2009
30 = Fournier et al. 2009
31 = Rizvi et al. 2009
32 = Kronström et al. 2009, 2011
33 = Cromer & Hilsenroth 2010
34 = McBride et al. 2010
35 = Steketee et al. 2011
36 = Lewis et al. 2012
37 = Renner et al. 2012
38 = Wolitzky-Taylor et al. 2012

Long-term therapies
39 = Jørgensen et al. 2000
40 = Puschner et al. 2004
41 = Solbakken et al. 2012

DIAGNOSIS (%)
Proportion of patients with:

AXIS I %   = Axis I disorders
  DEP %   = Depressive disorder
  ANX %   = Anxiety disorder
  COM I % = Comorbid Axis I disorders
AXIS II %   = Axis II disorders (i.e., personality disorders)
  COM II % = Comorbid Axis I and Axis II disorders

FOLLOW-UP
The length of follow-up after the end of treatment (in months).

MALES (%)
The percentage of male patients.

METHOD
1 = Correlation
2 = Modeling

N

Number of patients for which the patient characteristics are described (Table 6).
Number of patients used in the analysis (Supplementary tables 1A-F).

OUTCOME MEASURE

Treatment outcome was evaluated through psychiatric symptoms and functional ability.
These were classified into depressive symptoms, anxiety symptoms, and global
symptoms and functioning. The following measures for the evaluation of the symptoms
were used. Numbers in parentheses refer to the datasets in which they were used.

V1. Depressive symptoms

1. Self-rating Depression Scale (SDS) (1)
2. Raskin Depression Scale (2)
3. Hamilton Depression Rating Scale (HDRS) (2, 9, 10, 15, 16, 24, 25, 29, 30, 32)
4. Beck Depression Inventory (BDI) (6, 8, 9, 10, 12, 15, 16, 18, 21, 24, 34, 36)
5. Millon Clinical Multiaxial, Major depression (MCMI-Dep) (12)
6. Millon Clinical Multiaxial Inventory, Dysthymia (MCMI-Dys) (12)
7. Depression severity (based on BDI, HDRS, Inventory for Depressive
   Symptomatology – Self-report and Clinician-report) (23)
8. Montgormey-Åsberg Depression Rating Scale (MADRS) (26)

V2. Anxiety symptoms

1. a. Phobic Anxiety (Street subscale, self-report) (1)
   b. Phobic Anxiety (Street subscale, therapist-report) (1)
   c. Phobic Anxiety (Street subscale, independent observer -report) (1)
   d. Phobic Anxiety (Other phobias (average), self-report) (1)
   e. Phobic Anxiety (Other phobias (average), therapist-report) (1)
   f. Phobic Anxiety (Other phobias (average), independent observer -report) (1)
2. a. Phobic Avoidance (Street subscale, self-report) (1)
   b. Phobic Avoidance (Street subscale, therapist-report) (1)
   c. Phobic Avoidance (Street subscale, independent observer -report) (1)
   d. Phobic Avoidance (Other phobias (average), self-report) (1)
   e. Phobic Avoidance (Other phobias (average), therapist-report) (1)
   f. Phobic Avoidance (Other phobias (average), independent observer -report)
      (1)
3. a. Anxious mood (therapist-report) (1)
   b. Anxious mood (independent observer -report) (1)
4. Fear Survey Schedule (FSS) (1)
5. Beck Anxiety Inventory (BAI) (8)
6. Millon Clinical Multiaxial Inventory, Anxiety (MCMI-Anx) (12)
7. a. Mobility Inventory (agoraphobic avoidance when alone subscale (MI-AAL))
      (13)

      b. Mobility Inventory (agoraphobic avoidance when accompanied subscale (MI-AAC)) (13)

      c. Mobility Inventory (agoraphobic avoidance (MI-A), average score of MI-AAL (7a) and MI-AAC (7b)) (13, 28)

      d. Mobility Inventory (frequency of panic attacks subscale (MI-PF) (13, 28)

8. Physical panic symptoms (Nijmegen Hyperventilation List (NHL)) (13)
9. Compulsive behavior (Maudsley Obsessional Compulsive Inventory (MOCI)) (14)
10. Obsessive fear (Anxiety-Discomfort Scale; mean score of 5 patient, therapist and assessor ratings) (14, 28)
11. Trait Anxiety Scale (18)
12. Agoraphobic Cognitions Questionnaire (ACQ) (28)
13. a. Body Sensations Questionnaire (BSQ, fear subscale) (28)

      b. Body Sensations Questionnaire (BSQ, frequency subscale) (28)

14. CAPS (interview-based 17-item Clinician Administered PTSD Scale) (31)
15. Yale-Brown Obsessive Compulsive Scale (Y-BOCS) (35)
16. Mood and Anxiety Symptom Questionnaire, General Anxiety Subscale (MASQ-GA) (38)

V3. Global symptoms and functioning

1. Symptom Checklist-90, Global Severity Index (SCL-90-GSI) (3, 4, 7, 10, 12, 17, 18, 20, 22, 39, 41)
2. Symptom relief (5)
3. Global Assessment of Functioning Scale (GAF) (7, 33)
4. Global Assessment Scale (GAS) (10, 11, 22)
5. Improvement in symptoms (19)
6. Improvement in general functioning (19)
7. Mean Reliable Change Index (RCI) (based on BDI, SCL-90-R and diagnostic-specific measure depending on the primary presenting diagnosis: the Yale-Brown Obsessive Compulsive Scale, the Agoraphobic Conditions Questionnaire, the Social Phobia Scale and Social Interaction Anxiety Scale, the Why Worry Scale, the Structured Clinical Interview for DSM-III-R Axis II Disorder, Self-Report Version, or the psychosis subscale of the SCL-9-R) (27)
8. Brief Symptom Inventory, Global Severity Index (BSI-GSI) (33)
9. Symptom distress (45-item self-report Outcome Questionnaire, 25-item Symptom Distress subscale, OQ-SD) (40)

OUTCOME TYPE

1 = Binary post therapy value of outcome measure (i.e., remission around certain value)

2 = Continuous post therapy value of outcome measure

3 = Pre-Post therapy change in outcome measure

4 = Residual change score (i.e., Post score minus the change to be expected on the basis

of the regression of Pre score on Post score)

5 = Percentual improvement (i.e., Pre-Post / Pre)

6 = Reliable Change Index (RCI) (i.e., Pre-Post / standard error of the difference)

PREDICTOR MEASURE

The psychological suitability measures used to predict the treatment outcome were classified into 6 domains and 13 predictor categories. The following suitability measures were used. Numbers in parentheses refer to the datasets in which they were used.

D1. Nature of problems

P1. Focus

1. Focus (item of interview-based 6-item suitability selection criteria) (3)
2. Focality of complaint (item of interview-based 6-item selection criteria for Short-Term Anxiety Provoking Psychotherapy, STAPP, modified version 1) (5)
3. Psychodynamic explanation (item of interview-based 6-item selection criteria for Short-Term Anxiety Provoking Psychotherapy, STAPP, modified version 1) (5)
4. Circumscribed problems for treatment (factor based on interview-based 14-item selection criteria for Short-Term Anxiety Provoking Psychotherapy, STAPP, modified version 2) (11)
5. Identifiable focus (item of interview-based 39-item Checklist for Brief Psychodynamic Therapy) (19)
6. Focality (item of interview-based 10-item Suitability for Short-term Cognitive Therapy Rating Scale, SSCT) (27)

D2. Ego strength

P2. Modulation of affects

1. Ability to emote (item of interview-based 6-item suitability selection criteria) (3)
2. Able to recognize feelings (item of interview-based 39-item Checklist for Brief Psychodynamic Therapy) (19)
3. Accessibility of affects (factor based on interview-based 39-item Checklist for Brief Psychodynamic Therapy) (19)
4. Awareness and differentiation of emotions (item of interview-based 10-item Suitability for Short-term Cognitive Therapy Rating Scale, SSCT) (27)
5. Integrates affects (item of interview-based 9-item Capacity for Dynamic Process Scale, CDPS) (33)
6. Perceives affective aspects of problems (item of interview-based 9-item Capacity for Dynamic Process Scale, CDPS) (33)
7. Differentiates affects (item of interview-based 9-item Capacity for Dynamic Process Scale, CDPS) (33)

8. Affect (factor based on interview-based 9-item Capacity for Dynamic Process Scale, CDPS) (33)

9. Global affect consciousness (Affect Consciousness Interview, ACI) (41)

P3. Flexibility of interaction
1. Adequate contact with interviewer (item of interview-based 6-item selection criteria for Short-Term Anxiety Provoking Psychotherapy, STAPP, modified version 1) (5)
2. Involvement (factor based on interview-based 14-item selection criteria for Short-Term Anxiety Provoking Psychotherapy, STAPP, modified version 2) (11)
3. Alliance potential: in-session evidence (item of interview-based 10-item Suitability for Short-term Cognitive Therapy Rating Scale, SSCT) (27)
4. Collaborates therapeutically (item of interview-based 9-item Capacity for Dynamic Process Scale, CDPS) (33)

P4. Self-concept in relation to ego ideal
1. Developmental level of self-concept (interview-based Self-concept Rating Scale) (4)

D3. Self-observing capacity

P5. Trial interpretation
1. Ability to accept interpretations (item of interview-based 6-item suitability selection criteria) (3)

P6. Reflective ability
1. Psychological mindedness (item of interview-based 6-item suitability selection criteria) (3)
2. Psychological mindedness (score based on observer-rated Psychological Mindedness Assessment Procedure) (18)
3. Psychological mindedness (score based on self-reported 45-item Psychological Mindedness Scale, PMS) (32)
4. Manifests insight (item of interview-based 9-item Capacity for Dynamic Process Scale, CDPS) (33)
5. Appears introspective (item of interview-based 9-item Capacity for Dynamic Process Scale, CDPS) (33)
6. Manifests verbal fluency (item of interview-based 9-item Capacity for Dynamic Process Scale, CDPS) (33)
7. Insight (factor based on interview-based 9-item Capacity for Dynamic Process Scale, CDPS) (33)
8. Accessibility of automatic thoughts (item of interview-based 10-item Suitability for Short-Term Cognitive Therapy Rating Scale, SSCT) (27)

P7. Motivation

1. Motivation (item of interview-based 6-item suitability selection criteria) (3)
2. Motivation for psychotherapy (interview-based 38-item Motivation for Psychotherapy Scale) (4)
3. Motivation for change (item of interview-based 6-item selection criteria for Short-Term Anxiety Provoking Psychotherapy, STAPP, modified version 1) (5)
4. Motivation for change: Readiness-to-Change (RTC) score (self-reported 32-item University of Rhode Island Change Assessment Questionnaire, URICA) (35)
5. A. Readiness to change: Precontemplation score (subscale of self-reported 32 item Stages of Change Schedule, SOCS) (36)
   B. Readiness to change: Contemplation score (subscale of self-reported 32-item Stages of Change Schedule, SOCS) (36)
   C. Readiness to change: Action score (subscale of self-reported 32-item Stages of Change Schedule, SOCS) (36)
   D. Readiness to change: Maintenance score (subscale of self-reported 32-item Stages of Change Schedule, SOCS) (36)
6. Motivation (factor based on interview-based 14-item selection criteria for Short-Term Anxiety Provoking Psychotherapy, STAPP, modified version 2) (11)
7. Motivation for treatment (self-reported 12-item Nijmegen Motivation List, NML) (13, 14, 28)
8. Autonomous motivation (self-reported 6-item subscale of the Autonomous and Controlled Motivations for Treatment Questionnaire, ACMTQ) (24, 34)
9. Controlled motivation (self-reported 6-item subscale of the Autonomous and Controlled Motivations for Treatment Questionnaire, ACMTQ) (24, 34)
10. Expectation of improvement (Attitudes and expectations form: patient's global response) (10)
11. No problem ending therapy (item of interview-based 39-item Checklist for Brief Psychodynamic Therapy) (19)
12. Accepts problem is of his/her own doing (item of interview-based 39-item Checklist for Brief Psychodynamic Therapy) (19)
13. Positive attitude towards therapy (factor based on interview-based 39-item Checklist for Brief Psychodynamic Therapy) (19)
14. Acceptance of personal responsibility for change (item of interview-based 10-item Suitability for Short-term Cognitive Therapy Rating Scale, SSCT) (27)
15. Compatibility with cognitive rationale (item of interview-based 10-item Suitability for Short-term Cognitive Therapy Rating Scale, SSCT) (27)
16. Patient optimism/pessimism regarding therapy (item of interview-based 10-item Suitability for Short-term Cognitive Therapy Rating Scale, SSCT) (27)

D4. Intelligence

P8. Intelligence

1. Problem-solving capacity (item of interview-based 6-item selection criteria for Short-Term Anxiety Provoking Psychotherapy, STAPP, modified version 1) (5)
2. Problem-solving capacity (item of interview-based 14-item selection criteria for Short-Term Anxiety Provoking Psychotherapy, STAPP, modified version 2) (11)
3. Crystallized intelligence (self-reported 40-item WAIS-Clarke Vocabulary test, WAIS-C) (8)
4. Fluid intelligence (self-reported 40-item Abstractions subscale of the Shipley Institute of Living Scale) (8)
5. Intelligence (self-reported Shipley-Harford Living Scale, composed of 20-item verbal intelligence scale and 10-item analytic intelligence scale) (30)
6. Intelligence (self-reported 50-item Quick Test, QT) (31)

D5. Intrapsychic and interpersonal behavior

P9. Coping styles and cognitive skills

1. Dysfunctional attitudes (self-reported 40-item Dysfunctional Attitudes Scale, DAS) (6, 9, 15, 23, 26, 36)
2. Perfectionism (factor based on DAS) (10, 30)
3. Need for approval (factor based on DAS) (10, 30)
4. Dysfunctional attitudes, interpersonal items (of DAS) (16)
5. Dysfunctional attitudes, achievement items (of DAS) (16)
6. Achievement (subscale of DAS) (21)
7. Dependency (subscale of DAS) (21)
8. Self-control (subscale of DAS) (21)
9. Cognitive dysfunction (score based on DAS) (10)
10. Dysfunctional thinking (self-reported 30-item Automatic Thoughts Questionnaire, ATQ) (6)
11. Dysfunctional cognitive processing (self-reported 50-vignette Cognitive response test) (6)
12. Hopelessness (self-reported 20-item Hopelessness scale) (6, 30, 36)
13. Learned resourcefulness (self-reported 36-item Self-Control Schedule, SCS) (6, 9)
14. Frank's Mastery Scale (self-reported 17-item scale adapted from Seeman's Powerlessness Scale) (2)
15. Attributional style (self-reported 12-vignette Attributional Style Questionnaire, ASQ) (30)
16. Attributional style, composite index of failure (ASQ-F) (9, 23)
17. Attributional style, composite index of success (ASQ-S) (23)
18. Attributional style, interpersonal items (of ASQ) (16)

19. Attributional style, achievement items (of ASQ) (16)
20. Self-efficacy (self-reported 23-item Self-Efficacy Scale, SEF) (23)
21. Self-esteem (self-report10-item Rosenberg Self-esteem Scale) (30)

P10. Defense styles

1. Overall defensive functioning (rater-evaluated Defense Mechanism Rating Scale, DMRS-ODF) (22)
2. Overall defensive functioning (self-reported 88-item (Hersoug) / 42-item (Van) Defense Style Questionnaire, DSQ-ODF) (22, 29)
3. Immature defense style (factor based on DSQ) (29, 32)
4. Neurotic defense style (factor based on DSQ) (29, 32)
5. Mature defense style (factor based on DSQ) (29, 32)
6. Overall defensive functioning (observer-rated Developmental Profile, DP-ODF) (29)
7. Rivalry (score based on DP) (29)
8. Symbiosis (score based on DP) (29)
9. Resistance (score based on DP) (29)
10. Narcissistic (score based on DP) (29)
11. Fragmentation (score based on DP) (29)
12. Lack of structure (score based on DP) (29)
13. Generativity (score based on DP) (29)
14. Solidarity (score based on DP) (29)
15. Individuation (score based on DP) (29)
16. Security operations (item of interview-based10-item Suitability for Short-Term Cognitive Therapy Rating Scale, SSCT) (27)

P11: Personality traits

1. Therapist optimism: patient attractiveness and likability (item of interview-based 6-item suitability selection criteria) (3)
2. Creativity (factor based on interview-based 39-item Checklist for Brief Psychodynamic Therapy) (19)
3. Negative temperament (dimension of self-reported 375-item Schedule for Nonadaptive and Adaptive Personality, SNAP) (23)
4. Mistrust (dimension of SNAP) (23)
5. Self-harm (dimension of SNAP) (23)
6. Dependency (dimension of SNAP) (23)
7. Positive temperament (dimension of SNAP) (23)
8. Detachment (dimension of SNAP) (23)
9. Manipulativeness (dimension of SNAP) (23)
10. Aggression (dimension of SNAP) (23)
11. Eccentric perceptions (dimension of SNAP) (23)
12. Exhibitionism (dimension of SNAP) (23)
13. Entitlement (dimension of SNAP) (23)

14. Disinhibition (dimension of SNAP) (23)
15. Impulsivity (dimension of SNAP) (23)
16. Propriety (dimension of SNAP) (23)
17. Workaholism (dimension of SNAP) (23)
18. Novelty seeking (temperament dimension of self-reported 240-item Temperament and Character Inventory, TCI) (26, 32)
19. Harm avoidance (temperament dimension of TCI) (26, 32)
20. Reward dependency (temperament dimension of TCI) (26, 32)
21. Persistency (temperament dimension of TCI) (26, 32)
22. Self-directedness (character dimension of TCI) (26, 32)
23. Self-transcendence (character dimension of TCI) (26, 32)
24. Cooperativeness (character dimension of TCI) (26, 32)
25. Self-criticism (self-reported 66-item Depressive Experiences Questionnaire, DEQ) (24)
26. Dependency (self-reported 66-item Depressive Experiences Questionnaire, DEQ) (24)
27. Offers positive relationship (item of interview-based 9-item Capacity for Dynamic Process Scale, CDPS) (33)
28. Neuroticism (48-item self-report Maudsley Personality Inventory, MPI-N) (2)
29. Extraversion (48-item self-report Maudsley Personality Inventory, MPI-E) (2)
30. Neuroticism (self-reported 240-item NEO Personality Inventory – Revised, NEO-PI-R) (38)
31. Neuroticism (self-reported 60-item NEO-FFI, short form of the NEO-PI-R) (25)
32. Extraversion (self-reported 60-item NEO-FFI, short form of the NEO-PI-R) (25)
33. Openness (self-reported 60-item NEO-FFI, short form of the NEO-PI-R) (25)
34. Agreeableness (self-reported 60-item NEO-FFI, short form of the NEO-PI-R) (25)
35. Conscientiouness (self-reported 60-item NEO-FFI, short form of the NEO-PI-R) (25)
36. Anger-Hostility (self-reported 44-item State-Trait Anger Expression Inventory, STAXI) (31)
37. Guilt (self-reported 9-item Buss-Durkee Guilt Scale) (2)

P12. Interpersonal relationships
1. Interpersonal problems: problems with the significant partner (self-reported Marital Deprivation Scale, MDS, adapted version of Marital Attitude Evaluation Scale) (1)
2. Interpersonal problems: assertiveness (self-reported 48-item Adult Self-Expression Scale) (1)

3. Give-and-take relationship (item of interview-based 6-item selection criteria for Short-Term Anxiety Provoking Psychotherapy, STAPP, modified version 1) (5)

4. Quality of interpersonal relations (item of interview-based selection criteria for Short-Term Anxiety Provoking Psychotherapy (STAPP; modified version 2) (11)

5. Quality of object relations (interview-based Quality of Object Relations Scales, QORS) (18)

6. Stability of relationships (factor based on interview-based 39-item Checklist for Brief Psychodynamic Therapy) (19)

7. Transference manifestations present (item of interview-based 39-item Checklist for Brief Psychodynamic Therapy) (19)

8. Early parental relations (self-reported Structural Analysis of Social Behavior, SASB) (20)

9. Social Adjustment Scale – Self Report (self-reported 56-item SAS-SR) (23)

10. Social dysfunction (mean of 11 items in social and leisure activities subscale of SAS-SR) (10)

11. Feeling and satisfaction (subscale of modified 45-item SAS-SR) (26)

12. Interpersonal behavior (subscale of modified 45-item SAS-SR) (26)

13. Role performance (subscale of modified 45-item SAS-SR) (26)

14. Friction (subscale of modified 45-item SAS-SR) (26)

15. Interpersonal problems (self-reported 127-item Inventory of Interpersonal Problems, IIP) (23)

16. Domineering (subscale of circumplex model of IIP, IIP-C) (17)

17. Competitive (subscale of IIP-C) (17)

18. Overly Cold (subscale of IIP-C) (17)

19. Socially avoidant (subscale of IIP-C) (17)

20. Nonassertive (subscale of IIP-C) (17)

21. Exploitable (subscale of IIP-C) (17)

22. Overly nurturant (subscale of IIP-C) (17)

23. Intrusive (subscale of IIP-C) (17)

24. Overall affiliation (score based on IIP) (17, 40)

25. Overall dominance (score based on IIP) (17, 40)

26. Interpersonal distress (factor based on IIP-C) (37)

27. Agency (factor based on IIP-C) (37)

28. Communion (factor based on IIP-C) (37)

29. Overinvolved interpersonal style (Overinvolved scale of IIP) (21)

30. Underinvolved interpersonal style (Underinvolved scale of IIP) (21)

31. Secure interpersonal style (based on Interpersonal Scale, developed from IIP and the Attachment Vignettes) (21)

32. Avoidant interpersonal style (based on Interpersonal Scale, developed from IIP and the Attachment Vignettes) (21)

33. Ambivalent interpersonal style (based on Interpersonal Scale, developed from IIP and the Attachment Vignettes) (21)
34. Attachment anxiety (self-reported 30-item Relationship Scales Questionnaire, RSQ) (24)
35. Attachment avoidance (self-reported 30-item Relationship Scales Questionnaire, RSQ) (24)
36. Alliance potential: out-of-session evidence (item of interview-based 10-item Suitability for Short-term Cognitive Therapy Rating Scale, SSCT) (27)
37. Total Object-Relational Functioning (TORF score based on observer-rated Developmental Profile, DP) (29)
38. Rivalry (score based on DP) (29)
39. Symbiosis (score based on DP) (29)
40. Resistance (score based on DP) (29)
41. Narcissistic (score based on DP) (29)
42. Fragmentation (score based on DP) (29)
43. Lack of structure (score based on DP) (29)
44. Generativity (score based on DP) (29)
45. Solidarity (score based on DP) (29)
46. Individuation (score based on DP) (29)
47. Differentiates interpersonal events (item of interview-based 9-item Capacity for Dynamic Process Scale, CDPS) (33)
48. Relational factor (factor based on interview-based 9-item Capacity for Dynamic Process Scale, CDPS) (33)
49. Positive patient contribution to interpersonal process (based on self-reported 36-item Structural Analysis of Social behavior, SASB) (39)
50. Negative patient contribution to interpersonal process (based on SASB) (39)

D6. Overall suitability

P13. Overall suitability
1. Total predictor variable (TPV, sum of the values of the 6 items of interview-based suitability selection criteria) (3)
2. CDPS total score (average of the values of the 9 items of interview-based Capacity for Dynamic Process Scale, CDPS) (6, 33)
3. Mean SSCT (average of the values of the 9 (Safran) or 10 (Myhr) items of Suitability for Short-Term Cognitive Therapy Rating Scale, SSCT) (12, 27)
4. Suitability score (sum of the values of the 10 items of Dynamic Assessment Interview, DAI) (39)

PREDICTOR TYPE
1 = Single variable (i.e., single question, item of a scale)
2 = Summary variable (e.g., factor, summary score, scale)

RESULT
Statistically significant (p < 0.05) results / all results reported in the study.

SESSIONS
Mean number (range) of sessions attended if not otherwise mentioned.

STRENGTH OF ASSOCIATION
Minimum (Min), Median (Med), and Maximum (Max) values of correlation coefficients reported in the study. Positive values support the expected association while negative values do not.

NA = Not available (correlation coefficient)
NR = Not reported (strength of association)
^{0} (p ≥ 0.05)
* (p < 0.05)
** (p < 0.01)
*** (p < 0.001)

THERAPIST TRAINING
The level of therapist training and experience:

NT = Untrained (no formal training / work experience)
ST = Semi-trained (training ongoing or training finished but no work experience or under supervision)
T = Trained (formal training and work experience)

THERAPY / TREATMENT
Type of therapy or other treatment (in trials) received.

ACT = Acceptance and Commitment Therapy
BT = Behavioral Therapy
CBT = Cognitive-Behavioral Therapy
CT = Cognitive Therapy
CM = Clinical Management
ET = Experiential Therapy
IPT = Interpersonal Therapy
MED = Medication/Pharmacotherapy
NTC = Nonscheduled Treatment Control
PP = Psychodynamic Psychotherapy
PP(I) = Psychodynamic Psychotherapy (Interpretative)
PP(S) = Psychodynamic Psychotherapy (Supportive)
PLA = Placebo
WL = Waiting List

APPENDIX 3. INSTRUCTIONS FOR THE INTERVIEWER FOR THE ASSESSMENT OF SUITABILITY

Based on clinical observations from the semi-structured interview rate the score that best describes the person's functioning in each suitability variable.

Focus

1. The focus is clearly defined already during the first interview. The basic dynamic conflict is circumscribed and is reflected in the person's way of expressing the present difficulties. Often the person describes the focus in the form of a recollection of a past memory or as a mental image. When inquired about the background or reasons for the difficulties, he or she is able to produce material that corresponds to the idea of the focus formed by the interviewer.

2. The focus can be defined somewhat later, in the second or third interview. Otherwise the situation is as described above.

3. The focus and the dynamic conflict can be defined, but the person cannot provide experiential presentation. The person's thoughts about the background or reasons for the difficulties confirm the scope of the problems.

4. The determination of the focus is possible, but it remains global and unspecific. Theoretically, a hypothesis of the basic conflict can be constructed, but it is not expressed clearly in the person's narration. No clinical vignette can thus be observed. The person's conception of the problems is given as a general statement.

5. Determination of the focus is difficult. It remains global and can be stated as a dependency problem, separation-individuation problem, inability to deal with aggression, etc. Likewise, it can be related to the present problems without any background information of the dynamic problems. The person's conception of the problems is diffuse.

6. It is almost impossible to define a focus. There hardly exists a focus that could be worked with in short-term psychotherapy. The person's conception of the problems is diffuse.

7. The interviewer is unable to formulate any focus on the basis of a very diffuse presentation of the problems. The suitability of the person for psychotherapy in general is questionable.

Modulation of affects

1. The person describes him/herself in a balanced manner already during the first interview session. Wide array of both positive and negative affects are expressed. The person is in realistic contact with his or her emotional state and is capable of adequate control. There is no overload of affects and no difficulty in the continuity of the interview. The interviewer is able to identify with the affective expression of the person.

2. During the first interview, there is mild defensiveness; but by the next interview session, the person's expression and modulation of affects is similar to that described above.

3. In general, the person is not in good contact with his or her positive and/or negative affects, but manages to find it from time to time. There is some degree of defensiveness and increased control, but without major impact on the continuity of the interview.

4. There is significant defensiveness and shallowness of affects, which narrows the scope of the interview. Contact with the interviewer remains basically intact, however.

5. There are EITHER affective outbursts that make it difficult to continue the narration because of the emotional overflow OR significant defensiveness with a narrowing of the scope of narration, and temporary detachment from the contact.

6. The interview is characterized by affective outbursts and ruptures in the continuity OR the person is so defensive that no affective reaction can be seen. The presence of affects can be felt through counter-transference feelings, however.

7. The person is subject to a constant overflow of dysphoric affects, is agitated or emotionally aroused in a (hypo)manic state OR there is a rigid appearance with affective stupor or disaffectation. Contact with the interviewer is severely deteriorated and the interview cannot be carried out properly.

Flexibility of interaction

1. The interview can be carried out according to the protocol. The person is able to talk freely and answers the questions set out by the interviewer in an orderly, sequential way, remembering the different parts of the initial question. There is little need for additional questions, and the dialog with the interviewer is flexible and natural. The patient takes into consideration the time available.

2. The interview can be carried out according to the protocol, but some of the questions are not answered and the interviewer must repeat them. The dialog is natural. The interviewer may have to pay attention that the patient does not exceed the time available.

3. The interview can be carried out largely as described above (scores 1 and 2), but there are some themes in which the spontaneous narration is interrupted. The interviewer must repeatedly confront the person with unanswered issues and repeat the questions.

4. The person talks freely about the present situation but is restricted and needs additional questions when talking about earlier phases of life OR the person gives a diffuse description of the present situation and repeatedly returns to earlier life phases.

5. The interview does not proceed fluently. The interviewer is forced to ask separate themes one by one. When doing so, the patient is capable of producing material. Dialog is basically restricted and organized around the interviewer's separate questions.

6. The interview proceeds poorly. The person talks only about the present situation. When asked about earlier life phases, he or she answers with a few brief sentences. The course of life remains diffuse, and there is lack of continuity of the narration. Dialog is poor.

7. The interview does not proceed at all. The person answers all questions only partly and very briefly. The interviewer is unable to form a coherent view of the person. Dialog is very poor and consists of only questions and brief answers.

Self-concept in relation to ego ideal

1. The present self-concept is adequately balanced with ego ideal.

2. The person is prone to belittling him or herself, but ego ideal is attainable with present capacities and skills.

3. The person devalues him or herself. Ego ideal is realistic, but its attainment requires great effort and sacrifice and easily causes disappointments.

4. There are unrealistic and grandiose features in the person's self-concept and ego ideal. The person has a distorted way of seeing him or herself and aspires to aims that are beyond his or her capacities and skills. The grandiose features do not dominate the presentation, however.

5. The person belittles and devalues him or herself. There are grandiose features that disturb the balance between self-concept and ego ideal. The person suffers from repeated disappointments in certain life areas.

6. Self-concept and ego ideal are basically unrealistic and distort reality. The person lives in a world governed by grandiose conceptions and aims.

7. The person devalues him or herself and simultaneously demands the impossible and unattainable. The situation is characterized by stagnation and the experience of loss.

Trial interpretation

1. The material provided by the person makes it possible to frame a trial interpretation already during the first interview. His or her experiential reaction to it is clearly seen and leads to the elaboration of themes dealt with in the interview. The person also provides additional material. In the next session, the person spontaneously takes up the issue again and reports of having worked on it between sessions, and provides additional material.

2. As described above, except that it is not until the second or third interview session that the trial interpretation can be made.

3. The person reacts to the interpretation and returns to it during the next session. However, the additional material provided is not experiential, but intellectual. OR: The person provides some elaboration during the first interview session but does not take it up later.

4. The person's reaction to the trial interpretation leads to him or her providing additional material, but he or she does not return to it during the next session. OR: The interpretation can only be given during the third interview session.

5. The person may agree with the interpretation, but does not work with it any further and does not provide additional material.

6. The person ignores the interpretation. The material provided by the person leads the interviewer to make the interpretation on a general, explanatory level, which is poor and not well organized.

7. The interviewer is not able to make a trial interpretation, and therefore no reaction can be observed.

Reflective ability

1. The person describes him or herself in a comprehensive way, taking into consideration the time perspective. Memories from childhood and adolescence are elaborated and expressed vividly and are connected with presently-experienced affects. The person is able to link earlier experiences with the present. Introspective elaboration is organized around his or her own desires and hopes.

2. The person describes him or herself as above (score 1), but defensiveness limits the elaboration and material. However, he or she has emotional contact with memories, and there is less defensiveness after the first interview session.

3. The reflective presentation is limited due to defensiveness. Reflective ability may also be limited by a wide and comprehensive description, dominated by an increased amount of rationalization and intellectualization. The person is able to take into consideration the time perspective.

4. The person describes external facts and objects more than him or herself. He or she aims to protect him or herself in this way from reflecting on his or her own experiences. There is consideration of the time perspective, however.

5. Reflective ability is significantly restricted. The person concentrates only on the present or on the past. Efforts at linking these areas repeatedly lead to the person returning to either one or the other worlds of experience.

6. The person fails to reflect on experiences from either childhood or adolescence. Information given from these phases is restricted and mainly consists of external facts. He/she returns repeatedly to the description of the present situation.

7. There is no ability to reflect psychologically in an introspective fashion. He or she may instead offer somatic or genetic explanations. At best, psychological reflections are shallow and superficial.

Motivation

1. The person's motivation for psychotherapy is very good. He or she has sought psychotherapy at his or her own initiative after long-term consideration. He or she is reflective and willing to elaborate on the problems and to address them deeply and psychologically. He or she also understands the multitude of the problems.

2. The person's motivation for psychotherapy is good. He or she has sought psychotherapy at his or her own initiative, but the willingness to address the problems psychologically is less prominent than in score 1 described above.

3. The person's motivation for psychotherapy is good. Psychotherapy has been sought mainly based on the need aroused by the present symptoms. Expectations are therefore

related to symptomatic change, but the person also understands that there is a wider background to the problems.

4. Motivation for psychotherapy is quite good. Psychotherapy is sougth based on the need aroused by the present symptoms, and without them he or he would not have sought psychological help at all. Therapy is also based on someone else's recommendation or example. The person basically wants relief from the present situational problems.

5. Motivation is fair. The therapy is based on someone else's recommendation as much as on the person's own desire. The person is ambivalent regarding psychotherapy, and psychological introspection is largely missing. He or she is willing to try psychotherapy, but expectations are reserved. He or she would also consider medication, hypnosis, counselling etc.

6. Motivation is uncertain. Psychotherapy has been recommended for the person and his or her own attitude to it is largely ambivalent. He or she wants to be sure that the therapy would be helpful and finds it difficult to commit to the sacrifices of time and money required by the therapy. There is also a strong need to address the problems on his or her own. The person would rather use only medication or rely on receiving expert advice and suggestions.

7. The person is clearly not motivated to undergo psychotherapy.