



Measuring Verbal Psychotherapeutic Techniques—A Systematic Review of Intervention Characteristics and Measures

Antje Gumz^{1,2*}, Barbara Treese³, Christopher Marx¹, Bernhard Strauss³ and Hanna Wendt^{2†}

¹ Berlin University of Psychology, Berlin, Germany, ² Department of Psychosomatic Medicine and Psychotherapy, University Medical Center Hamburg-Eppendorf, Hamburg, Germany, ³ Institute of Psychosocial Medicine and Psychotherapy, University Hospital Jena, Jena, Germany

OPEN ACCESS

Edited by:

Wolfgang Tschacher,
University of Bern, Switzerland

Reviewed by:

Alemka Tomicic,
Universidad Diego Portales, Chile
Agnes Von Wyl,
Zurich University of Applied Sciences,
Switzerland
Louis Castonguay,
Pennsylvania State University, USA

*Correspondence:

Antje Gumz
a.gumz@
psychologische-hochschule.de

† Present Address:

Hanna Wendt,
Schön Klinik Bad Bramstedt,
Bad Bramstedt, Germany

Specialty section:

This article was submitted to
Psychology for Clinical Settings,
a section of the journal
Frontiers in Psychology

Received: 21 July 2015

Accepted: 22 October 2015

Published: 10 November 2015

Citation:

Gumz A, Treese B, Marx C, Strauss B
and Wendt H (2015) Measuring Verbal
Psychotherapeutic Techniques—A
Systematic Review of Intervention
Characteristics and Measures.
Front. Psychol. 6:1705.
doi: 10.3389/fpsyg.2015.01705

Language is one of the most important “tools” of psychotherapists. The working mechanisms of verbal therapeutic techniques, however, are still marginally understood. In part, this is due to the lack of a generally acknowledged typology as well as a gold standard for the assessment of verbal techniques, which limits the possibility of conducting studies focusing this topic. The present study reviews measures used in clinical research which assess directly observable dimensions of verbal interventions in a reliable manner. All measures were evaluated with respect to their theoretical foundation, research goals, assessment modes, and various psychometric properties. A systematic search in databases (PubMed, PsycInfo, PsycArticles, PSYINDEX, Web of Science, Embase) followed by an additional “snowballing” search covering the years 1940–2013 yielded $n = 179$ publications eligible for review. Within these publications, 34 measures were identified showing great heterogeneity regarding the aspects under study. Only two measures reached the highest psychometric standards and can be recommended for clinical use without any reservation. Central problems include deficiencies in the systematization of techniques as well as their partly ambiguous and inconsistent definitions. To promote this field of research, it will be important to achieve a consensus concerning the terminology, conceptions and measures of verbal interventions.

Keywords: verbal intervention, technique, skills, measure, psychometric property, systematic review

INTRODUCTION

The general efficacy of psychotherapy is well established (e.g., Lambert, 2013) based on a variety of reliable and valid measures. Numerous studies demonstrated that psychotherapeutic treatment using different change theories and related to diagnoses is associated with an improvement of symptoms, quality of life and psychological, social and occupational functioning (e.g., Beutler, 2009; Norcross, 2011; Wampold and Imel, 2015). Moreover, there seem to be only minor significant differences of therapeutic effects in different psychotherapy schools (Castonguay and Beutler, 2006; Barber et al., 2013; Wampold and Imel, 2015). However, other findings indicate that there is still a considerable amount of patients who do not benefit from psychotherapy (Lambert, 2013), or even get worse and that the sustainability of therapy effects is limited in certain patient populations (Barlow, 2010).

These findings call for increased research efforts in order to improve mental health care. Accordingly, the focus of psychotherapy research has increasingly shifted from the question “if” psychotherapy is effective to the question what works and how it works (Castonguay, 2013).

The Active Ingredients of Psychotherapy

What are the active ingredients of psychotherapy? In order to answer this question it is important to decide which variables need to be considered in respective research agendas. Psychotherapy processes can be described with regard to different characteristics (e.g., Orlinsky and Howard, 1987). An important topic in this respect is whether the effect of psychotherapeutic interventions is mainly due to common therapeutic factors or to therapeutic techniques.

Common factors are implicitly relevant in any therapeutic interaction. They are not explicitly anchored in the treatment models of the different schools of psychotherapy, nor are they considered in the treatment models for specific mental disorders (Tschacher et al., 2014). In their Taxonomy Project Tschacher et al. (2014) address the topic of definition and conceptualization of common factors of psychotherapy and delineate different terminological problems and logical inconsistencies. Based on a comprehensive literature search the authors specified all constructs discussed as non-specific or common factors of therapeutic change in psychotherapy research literature by at least two authors. The resulting list comprises 22 common factors (Pfammatter and Tschacher, 2012; Tschacher et al., 2014). These factors include, for example, therapeutic alliance, readiness to change, insight or cognitive restructuring (Connolly Gibbons et al., 2009; Tschacher et al., 2014). In accordance with the Generic Model of Psychotherapy (Orlinsky and Howard, 1987) common factors relate to various aspects of the therapeutic process: Interpersonal (e.g., therapeutic alliance), intrapersonal (e.g., instillation of hope), or clinical aspects (e.g., affective catharsis).

In many cases, techniques which are considered to belong to a particular psychotherapy method, and common factors have been treated as alternatives in explaining therapeutic progress. In other words, it was considered that *either* specific therapeutic interventions *or* common factors can explain the effects of psychotherapy. This dichotomy, however, needs to be questioned. Techniques and common factors cannot be considered independently. Common factors evolve in the context of the therapeutic relationship, which in turn is influenced to a large extent by the techniques. Having this in mind, common factors can be regarded as active principles implemented by the therapist’s specific techniques (Castonguay and Beutler, 2006; Pfammatter and Tschacher, 2012; Gumz et al., 2013).

Knowledge about the common factors provides little guidance for the therapist for what can actually be done in the therapy sessions. Empirical knowledge regarding the therapeutic techniques, i.e., regarding the question what therapists mostly do in successful sessions or successful therapies in general, are of great value to psychotherapy, as they allow recommendations of specific practice. Differentiation of therapist behaviors is crucial for specifying the extent to which techniques actually differ

and how this ultimately affects treatment outcome (Bergin and Strupp, 2009). Therefore, it is an important desideratum to analyze therapeutic techniques on the level of single sessions or single units of therapeutic interaction to give profound answers to the question of the how of therapeutic effectiveness (Margison et al., 2000; Mergenthaler, 2015).

What is Specifically Meant by Technique?

To date, there is no general consensus regarding the question what is meant by the term “technique.” Is this what the therapist does or says? Obviously not. The term can also refer to therapeutic attitudes, implicit theories or even patient behavior typical to a therapy method, e.g., warmth and empathy (Bergin and Strupp, 2009), therapeutic abstinence, neutrality, free association technique or regression (Gumz et al., 2014; Tschacher et al., 2014). Some techniques refer to broader descriptions of procedures or settings, e.g., role playing, sculpture work, reflecting team technique and hypnosis. In other cases, “technique” refers to detailed descriptions of single verbal statements such as the verbalization of emotional reactions or transference interpretations (Gumz et al., 2014; Tschacher et al., 2014).

Bergin and Strupp (2009) pointed out that the repertoire of the contemporary psychotherapist includes an impressive list of techniques, which often are employed in combination but more or less intuitively and unsystematically. Schools of therapy differ in their relative emphases upon particular techniques.

In our view, it is possible to draw a distinction between specific and common techniques. The term “specific” means that a certain technique is unique for a particular psychotherapy method. For example, exposure with response prevention is considered as a typical technique of cognitive-behavioral therapists, transference interpretation as a typical technique for psychodynamic therapists and circular questions as a typical technique for systemic therapists. Common techniques are those that are not specific for a respective therapy method. For example, the technique of exploration cannot be related to a specific therapy method. Another example is the verbalization of emotional reactions which is considered as typical technique of humanistic therapists, while psychodynamic therapists use this technique as well (Gumz et al., 2014).

The term “specific” might also mean that a certain technique may be tailored to the treatment of a specific psychiatric disorder. Such a tailored technique is, for example, planning of pleasant activities in the treatment of depressive disorders or keeping a food diary in the treatment of eating disorders (Pfammatter and Tschacher, 2012; Tschacher et al., 2014).

All in all, therapeutic techniques are characterized by numerous very different features, which makes it difficult to establish definite criteria for their assessment (Gumz et al., 2014). One approach to reduce complexity is to focus on single verbal techniques as a first step. On this level, however, an orientation is not easy as well, as there is no generally acknowledged typology of verbal techniques. Types of verbal techniques are described inconsistently with regard to the number and kind of techniques of a certain therapeutic method. Moreover, definitions of categories with the same label partly do not match, or vice

versa, identical aspects have differing designations (Brumberg and Gumz, 2012; Gumz et al., 2014).

This can be illustrated using the example of clarification which was analyzed amongst other categories within the framework of a systematic investigation of definitions of psychodynamic intervention techniques in the theoretical literature (Gumz et al., 2014). In most, but not in all of the analyzed sources, clarification was specified as a type of intervention technique of high importance. Different authors consistently specified that the therapeutic aim and a core characteristic of a clarification is to foster the understanding of a phenomenon. This aim is also described as a core characteristic of the concept interpretation. The formal techniques through which this aim is supposed to be achieved, are heterogeneous. Thus, the clarification was described as: (a) The patient is asked to specify and to associate or (b) the patient is asked to describe a phenomenon in more detail or (c) the therapist identifies recurrent topics or themes or (d) the therapist inquires about the patient's feelings (e.g., with respect to related associations or parallels) or (e) the therapist describes the effect of the patient's behavior on the therapist or, finally, (f) the therapist rephrases the central ideas of the patient's statement or summarizes these.

In order to clarify what is specifically meant by "therapeutic techniques," we suggest to systematize on four hierarchical levels:

Level (1) First of all, techniques referring to the therapeutic dialogue need to be distinguished from techniques referring to broader descriptions of procedures or settings or physical exercises, e.g., role playing, sculpture work, reflecting team technique and hypnosis, keeping a food diary in the treatment of eating disorders, or telling the patient to do a breathing exercise.

Level (2) Within the group of techniques referring to the therapeutic dialogue, verbal, and non-verbal techniques can be distinguished. We define verbal techniques as the verbal utterances of the therapist within the therapeutic dialogue. Non-verbal techniques (or behavior, communication) are closely intertwined with verbal information and prosodic features (Madonik, 2001; Pawelczyk, 2011; e.g., mimic signals, affective expressions, movement patterns, see Geißler, 2005).

Level (3) Verbal techniques contain directly observable and latent features. *Directly observable features* relate to the semantic content ("what is said") of therapeutic utterances or other semantic units (words, sentences, longer segments) which can be rated based on session transcripts or audio recordings. *Latent features* characterize the implicit pragmatic content of the utterances ("what is implicated" or "what is meant"). In order to evaluate latent features of verbal techniques a higher degree of subjective inference is necessary, compared to directly observable features. If a whole session is rated (global coding), the semantic or implicit pragmatic content of all therapeutic utterances is usually aggregated. In consequence, global coding methods generally involve a higher degree of subjective inference because larger amounts of data need to be cognitively aggregated and there are no obligatory rules regarding this aggregation (Heaton et al., 1995). Among the latent features there are, e.g., functional characteristics (e.g., therapist's intentions such as "directing the dialog," "speaking kindly," "controlling the affects" or therapeutic attitudes like abstinence, neutrality) or qualitative

characteristics (e.g., therapist's empathy, warmth or competence, internal coherence of an intervention).

Level (4) Regarding directly observable features we suggest to differentiate three basic characteristics: 1. *Form* (i.e., the formal and structural manner in which the therapist responds to the patient's experience, behavior or statement, e.g. "closed question," "interpreting," "paraphrase"), 2. *Temporal focus* (i.e., the period of time to which the intervention refers, e.g., "childhood," "present"), and 3. *Thematic content* (i.e., the topic of the intervention, e.g., "therapist," "mother," "defenses").

Research on Verbal Therapeutic Techniques

Within psychotherapy process research, varying instruments are used to measure therapeutic techniques, each having its own focus, capturing different features, and showing different psychometric characteristics. The findings of empirical studies are highly variable and often inconsistent, making it difficult to draw firm conclusions regarding the nature, processes, and effects of verbal techniques. There is an overwhelming amount of results concerning different techniques from various psychotherapy schools, naturalistic or manualized, gained by means of observational studies or interventional studies and on the basis of very different sample sizes. These studies examined the association of techniques with different factors such as symptom change, therapeutic relationship, therapeutic gains in single sessions or the relationship of specific techniques with therapy- or therapist-related variables. The diversity of these studies is difficult to systematize.

It can be assumed that all these difficulties contributed to the fact that, so far, there is little systematic knowledge regarding the efficacy of therapeutic techniques. Reviews on the subject are either concerned with selected types of techniques (e.g., transference interpretation; Høglend, 2004; Brumberg and Gumz, 2012), methodologically flawed (e.g., Blagys and Hilsenroth, 2000, 2002), or out-dated (Elliott et al., 1987).

An important first step to advance this field of research should be to establish a more solid theoretical as well as methodological framework for the measurement of verbal techniques. As a starting point we suggest to differentiate four basic features of verbal techniques as explained in the previous section: 1. *Form* (observable characteristic), 2. *Temporal focus* (observable characteristic), 3. *Thematic content* (observable characteristic), and 4. *Latent characteristics*. Such a typology can be helpful to review and structure the amount of existing theoretical articles and empirical results regarding categories of verbal therapeutic techniques.

The Current Review

Systematic knowledge regarding the question what kind of verbal techniques a therapist should or should not apply in his or her therapeutic work is of major practical importance. Our review attempts an initial clarification of the state of affairs by addressing the question which instruments are available for the assessment of verbal techniques in psychotherapy, which types of techniques are assessed and which psychometric properties the measures have.

We will restrict our review to the directly observable features of verbal techniques (see Section What is Specifically Meant by Technique?). Verbal techniques are the cornerstone of most psychotherapeutic methods. Language is the basic medium through which new information is conveyed, and it is one of the most important “tools” of the therapist (Gumz et al., 2014). Verbal techniques can, in principle, be examined beyond the differences between therapeutic schools. Focusing on verbal techniques, we will concentrate on measures used to assess directly observable dimensions. This means that we will restrict the review on measures which assess the formal, temporal, and thematic dimensions. Measures that exclusively address latent characteristics of therapeutic techniques were not considered in this review, as this would have potentiated the complexity of the research agenda and in the light of the long-term objective to systematically analyze techniques in order to be able to establish the state of the art in this research area. By focusing on explicit features of verbal techniques it is possible to clarify major facets of the research agenda. The next step will then be to extend the focus of research on latent features.

Moreover, an important criterion was that the reviewed measures report reliability.

We hope to contribute to a more consistent understanding of the subject. A more advanced systematization will help researchers to evaluate results as well as to choose existing instruments or to develop new instruments. Moreover, there is a practical benefit of our review: Only if the categories of a measure can be reliably registered by different individuals (interrater reliability), they can later be correctly trained and applied by other clinicians.

Our specific research questions were:

1. What are the characteristics of existing reliable measures assessing the formal, temporal and thematic dimensions of verbal therapeutic techniques?
 - a. What is the theoretical basis for the measures?
 - b. How have the measures been developed and which research questions were addressed?
 - c. What are the scales, categories, or items underlying the measures?
 - d. Which assessment mode is employed and which rating perspectives are considered?
2. What are the psychometric properties of the measures?

METHODS

Eligibility Criteria

Studies were eligible for inclusion if they met the following criteria:

- a. The study dealt with verbal therapeutic techniques within a psychotherapeutic setting.
- b. Verbal therapeutic techniques were assessed at least partially with regard to formal, temporal or thematic features.
- c. The study was based on ratings of specific psychotherapy sessions.

- d. The study referred to individual outpatient psychotherapy with adults.
- e. The study was published in a peer-reviewed journal.
- f. The study was published in English or German.
- g. The measure was developed with regard to well-established methods assuring the interrater reliability of the assessment of verbal techniques in order to establish its general applicability, i.e., Intraclass Coefficient Correlations (ICC; Shrout and Fleiss, 1979), (weighted) kappa-values (Cohen, 1960, 1968; Fleiss, 1971; Light, 1971), Finn's r (Finn, 1970, 1972), or Pearson's product-moment correlation coefficient. The procedure of reliability assessment and the reliability values need to be reported in a study published in English or German in a peer-reviewed journal.

Search Procedure and Study Selection Process

A systematic database search (Pubmed, PsycInfo, PsycArticles, PSYINDEX, Web of science, and Embase) for the years from 1940 to 2013 was conducted by one of the authors (B.T.) using the following search terms, their combination and truncation: psychotherapy, psychotherapeutic process, process research, measure, scale, rating, instrument, intervention, therapeutic technique, coding. The resulting records were judged by B.T. regarding titles and abstracts considering the above mentioned eligibility criteria. The 162 resulting studies were judged for eligibility by all authors. In a next step, further relevant articles were identified (by B.T. and H.W.) by screening references of the included articles as well as the excluded reviews (“snowballing,” c.f. Greenhalgh and Peacock, 2005) and by searching identified authors' names and measures (hand search). Additionally, the authors of all measures were contacted to collect coding manuals and further relevant information (non-response to the request was not a criterion for exclusion).

Psychometric Evaluation of Measures

Each measure was evaluated regarding objectivity, reliability, and validity. For each of these criteria, an “A” means that there are no limitations, a “B” indicates some restrictions, and “C” points to severe deficits (see **Figure 1** for evaluation criteria). An existing coding manual with clear and unambiguous item/category definitions enables the user to conduct objective ratings. Therefore, we assumed high objectivity when a detailed coding manual was published or available from the authors. Since all included measures report reliability, this aspect is differentiated into two levels with ICC values being the best because they are in line with the highest statistical standards (Bartko and Carpenter, 1975; Tinsley and Weiss, 1975).

RESULTS

Study Selection

Based on the systematic literature research, 1568 records were found. After screening the titles and abstracts, 162 records were selected for assessment of eligibility. Ninety-nine records had to be excluded: 31 reviews and 26 records were not peer-reviewed (11 book chapters, 15 dissertations), 23 records did not

| critterion | objectivity | reliability | validity |
|------------|---------------------------------------------------------------------|-------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------|
| A | a detailed coding manual is published or available from the authors | demonstrated by good ICC (2,k) values | demonstrated by several analyses regarding different aspects of validity |
| B | a coding manual exists as an unpublished manuscript | demonstrated by good kappa, Fleiss' r, or Pearson product-moment correlations | demonstrated by a few analyses regarding one or two aspects of validity (face validity is not admitted) |
| C | no reference to any coding manual | | demonstration of validity failed OR no (sufficient) information about validity |

FIGURE 1 | Criteria for evaluation of the included measures.

refer to verbal therapeutic techniques, but to other aspects of psychotherapy. Further six records had to be excluded, because the measure did not refer to the intervention form or content characteristics of verbal techniques. Five further records were excluded because they were not used for rating therapy sessions but the general activity of therapists in psychotherapy. Seven measures had to be excluded because of insufficient interrater reliability (e.g., percentages of agreements) or because reliability was not reported at all (Adler and Enelow, 1966; Holzman and Forman, 1966; Karl and Abeles, 1969; Gedo and Schaffer, 1989; Winston et al., 1991; Bucci and Maski, 2007; Hepner et al., 2010). Only one record was published in another language than English or German so that it could not be included. The following snowballing research revealed 116 further relevant records (see **Figure 2** for study selection process). In total, 179 articles were included in the present systematic review. They comprise 168 peer-reviewed articles and 11 published coding manuals or scale descriptions and comprise a total of 34 measures. These are presented in **Tables 1–5** in an alphabetical order, separated according to the theoretical orientation and the assessment mode (global measures referring to complete therapy sessions vs. microanalytic measures referring to therapists' single utterances). To maintain a good readability we will refer to the respective measure using its acronym (e.g., YACS for the Yale Adherence and Competence Scale, see **Tables 1–5** for the references of respective acronyms).

Theoretical Orientation

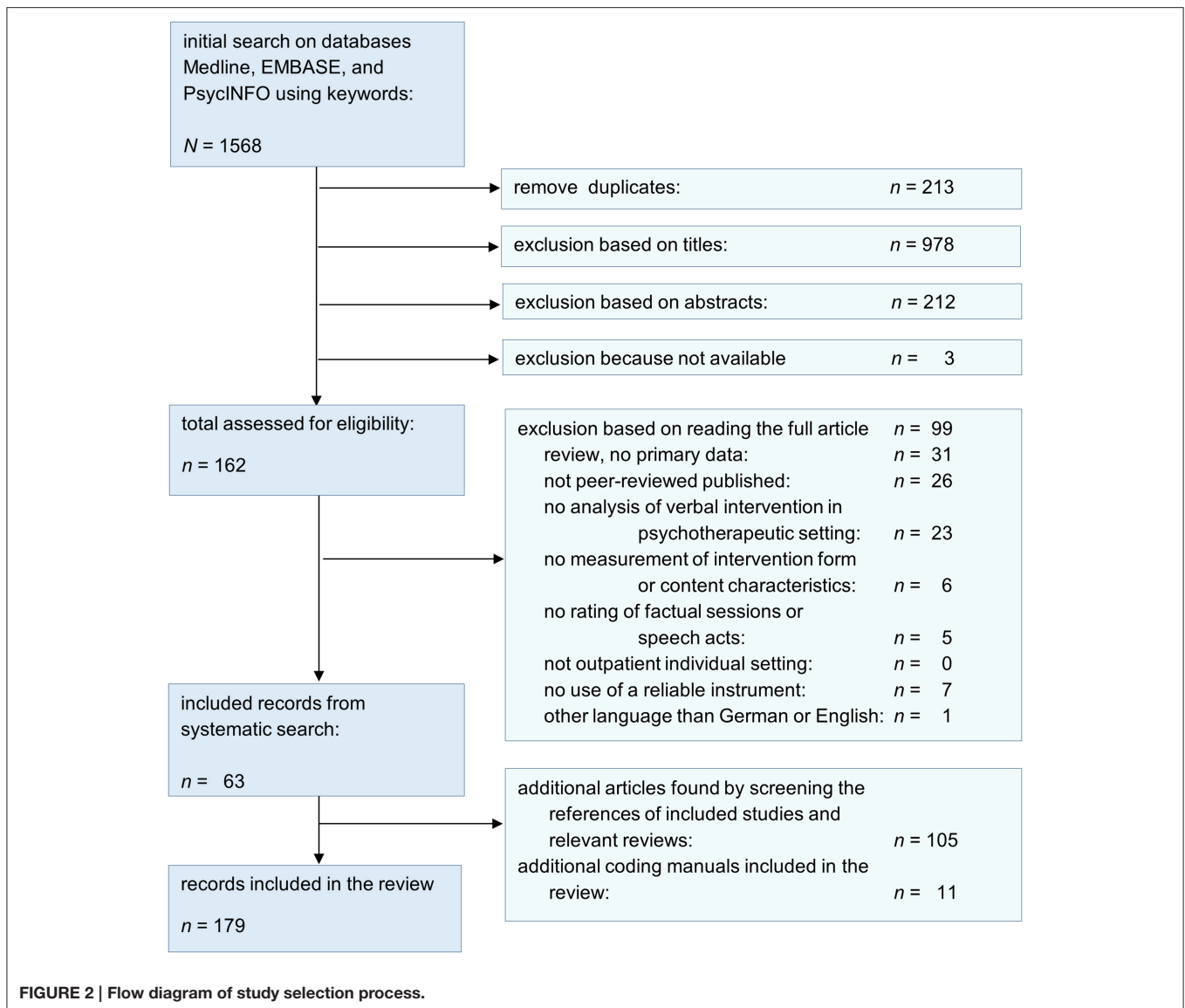
Eighteen of the included measures refer to a specific theoretical orientation. Most of these are derived from psychodynamic therapy (see **Table 1** for global and **Table 2** for microanalytic measures). Within these, two measures are specially developed for use in analytic therapy, two for supportive-expressive therapy and one for time-limited psychodynamic therapy, the other seven

measures can be generally used in dynamic psychotherapy. Only three measures refer to cognitive-behavioral orientated therapy or to a specific setting (drug counseling, treatment for substance use disorders, motivational interview), respectively (see **Table 3**). Sixteen measures were classified as “pantheoretical,” because they refer to several different theoretical orientations (see **Table 4** for global and **Table 5** for microanalytic measures).

Scale Development and Research Questions

All measures have been developed using a top down approach based on theoretical considerations, therapy manuals, literature research, clinical experience, and expert discussion as well as already existing scales. Three measures (APS, see **Table 2**; YACS, see **Table 3**; and PQS, see **Table 4**) included bottom-up analyses comprising analyses of therapist's utterances and audio- or videotaped therapy sessions among other strategies within the scale development process but none of the measures was developed exclusively in a bottom-up procedure (see **Tables 1–5**, column 3).

The measures have been developed to address various research questions. These can be classified into five groups: 1. Six measures aim to assess the adherence and competence of the therapist regarding specific therapeutic orientations (e.g., ACS-SEC, see **Table 1**; CTS-R, see **Table 3**; SPRS, see **Table 4**; YACS, see **Table 3**). 2. Among the pantheoretical measures, five measures aim to distinguish between two or more therapeutic orientations (e.g., CPPS, see **Table 4**; CSPRS, see **Table 4**; CSTF, see **Table 5**; SATC, see **Table 5**). 3. Two measures aim to investigate microprocesses in psychotherapy such as patient therapist interaction (e.g., PIRS, see **Table 2**; and CIP, see **Table 3**), or to analyze the relationship between therapeutic techniques and therapy outcome (four measures, i.e., APS, see **Table 2**; CAPS, see **Table 1**; MIT, see **Table 2**; PIC see **Table 2**). 4. Eleven measures



try to assess all possible verbal techniques in general (e.g., ITS, see **Table 5**; PQS, see **Table 4**) or within a therapeutic orientation (e.g., Coding of therapist statement, see **Table 2**; CTACS, see **Table 3**). 5. Finally, six measures focus on very specific aspects of psychotherapy, e.g., techniques used within the psychodynamic therapy of patients with borderline personality disorder (TVII, see **Table 1**), or focusing theoretically important constructs such as “interpretations” (TIRS, see **Table 2**).

Description of Subscales, Categories, and Items

According to the research questions, the measures assess verbal therapeutic techniques on various differentiation levels. The number of categories ranges from four categories without further differentiation (Coding of therapist statement; see **Table 2**; Connolly et al., 1998) to 100 items without further categorization (PQS; see **Table 4**; Jones, 1985) or 96 items, categorized into eight

scales with several subscales (CSPRS; see **Table 4**; Hollon et al., unpublished work, cited from Hill et al., 1992; see **Figure 3** for details). Apart from this differentiation level, the focus is partly limited to only one or two aspects of verbal techniques. The MIT (see **Table 2**) for example differentiates between “interpretation” vs. “non-interpretation.” The ISTS (see **Table 1**) and the ITS (see **Table 5**) distinguish the scales “supportive techniques” from “interpretative techniques” which include further items, respectively. The category or item that is assessed by the majority of measures is “interpretation.”

All measures assess techniques concerning their form (e.g., “question,” “interpretation,” “agenda setting”). Moreover, ten measures refer to the content of an intervention. Among those which address the thematic content or the temporal focus are for example the PIC (see **Table 2**; four content codes: defenses, anxieties, impulse/feeling, no content), the TAS (see **Table 1**; “future plans of patient”), and Response modes coding system

TABLE 1 | Global measures referring to psychodynamic therapy.

| Instrument source | Basis Theoretical orientation Objective within the development | Short description | Assessment mode | | Valuation | | |
|------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------|-------------|-------------|----------|
| | | | Raters | Definition of scales | Objectivity | Reliability | Validity |
| Adherence/Competence Scale for SE for Cocaine Dependence developed within the NIDA-CCTS (ACS-SEC; Barber et al., 1997) | Luborsky's manual for supportive-expressive treatment (Luborsky, 1984) manual of Mark and Faude (1995); own work (PACS-SE, Barber and Crits-Christoph, 1996) SE To assess the frequency of a technique and the competence with which it is implemented | 82 items, 3 subscales: 1. supportive (13 items) 2. expressive (31) 3. cocaine abuse (11) 17 additional items from the original scale PACS-SE | m (7-point Likert scale), session O | No coding manual published, description of scales within (Barber et al., 1997) | C | A | B |
| Columbia Analytic Process Scale (CAPS; Vaughan et al., 1997) | Review of clinical literature, meetings with another study group, TIRS (Bordin, 1966; Piper et al., 1987) AP Specifically designed for use in an outcome study | 3 parts: 1. free association (level I, II or III) 2. interpretation vs. non-interpretation (formal intervention, information providing, questions, advice) 3. working through (insight, understanding regarding an observation about the self, fantasy, transference, genetic history) | d, session O | Short coding manual within (Vaughan et al., 1997) | C | B | C |
| Interpretive and Supportive Technique Scale (ISTS; Ogrodniczuk and Piper, 1999) | Interpretive/supportive treatment manuals, literature review, clinical experience PDT Measures interpretive and supportive features of technique for a broad range of dynamically oriented psychotherapies | 14 items, 2 scales: 1. supportive techniques: gratify, noninterpretive interventions, guidance, problem solving, explanations, praise, personal information 2. interpretative techniques: pressure, uncomfortable emotions, interpretations, impression of therapist, linking, patient-therapist relationship, impression of others | m (5-point Likert scale), ratings of single interventions complete the overall rating on session-level O | Coding manual available from the authors | A | A | B |
| Therapist Action Scale (TAS) parallel form: Patient Action Scale (PAS) (Hoyt et al., 1981) | Previous therapist-activity rating systems, process-outcome studies, theoretical and clinical background PDT To assess the emphasis of specific actions of therapist and patient | 25 items, no subscales e.g., self-concept of patient, expression of feelings, prescription/advise given, future plans of patient | d, m, session O | No coding manual published; categories within (Hoyt et al., 1981) | C | B | B |
| Therapist Verbal Interventions Inventory (TVI; Koenigsberg et al., 1985) | Review of instruments, interventions described in literature, clinical experience PDT To measure the range of interventions used by psychodynamically oriented therapists in therapy of patients with borderline personality disorder | 35 items, 20 categories (further specified into 4 subcategories): provides factual information, seeks clarification, confronts, interprets) e.g., gives advice, offers sympathy, focus on external reality | d, 15-min segment O | Coding manual exists as an unpublished manuscript | B | B | C |

(Continued)

TABLE 1 | Continued

| Instrument source | Basis Theoretical orientation Objective within the development | Short description | Assessment mode Raters Definition of scales | Valuation | | |
|---------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------|-------------|-------------|----------|
| | | | | Objectivity | Reliability | Validity |
| Vanderbilt Therapeutic Strategies Scale (VTSS; Butler et al., 1995) | Manual for time-limited PDT, literature on psychodynamic technique Time-limited PDT Adherence scale to differentiate between experienced and unexperienced therapists | 21 items, 2 subscales: 1. interviewing style (IS, 12 items) 2. TDLP-specific strategies (SS, 9 items) | m (5-point Likert scale), session O Coding manual exists as unpublished manuscript | B | A | B |

m, metric; d, dichotomous; O, observer; SE, supportive-expressive dynamic psychotherapy; AP, analytic psychotherapy; PACS-SE, Penn Adherence-Competence Scale for Supportive-Expressive therapy; PDT, psychodynamic therapy; TIRS, Therapist Intervention Rating System; TDLP, Time-limited-dynamic psychotherapy; NIDA-CCTS, National Institute on Drug Abuse-Collaborative Cocaine Treatment Study. Evaluation criteria: A, no limitations; B, some restrictions; C, severe deficits (see text of the manuscript for details).

(see Table 5; three categories: statement categories, time frames, and person code). But 12 measures also refer to the latent dimension, because they include scales or items which address the “climate or atmosphere” (PQS, see Table 4) or the “Therapist’s warmth and friendliness” (VPPS, see Table 4). Some measures even mix the formal and latent dimension on item levels (e.g., VTSS, see Table 1: “Therapist shows evidence of listening respectively,” ISTS, see Table 1: “gratify the patient, i.e., make the patient feel good rather than anxious in the session”).

Assessment Mode and Rating Perspective

The assessment mode of the included measures can be generally classified into global vs. microanalytic. Global measures refer to whole therapy sessions or larger segments of sessions. Microanalytic measures refer to more fine-grained units like utterances or sentences. Furthermore, judged units can be assessed on metric vs. dichotomous scales and can be coded as manifold vs. mutually exclusive.

Within the theoretical orientations, the amount of global and microanalytic measures is approximately counterbalanced. Global measures generally use metric scales, microanalytic measures prefer dichotomous items which are mutually exclusive. The only measures that assess verbal therapeutic techniques in a metric way on a microanalytic level are: 1. the APS (where ratings refer to segments of session so that it is not purely microanalytic, see Table 2) and one scale (“degree of initiative”) of the SATC (see Table 5), which does not belong to the observable formal or content dimension of verbal techniques. All in all, no measure purely measures verbal techniques in a metric way on a microanalytic level.

The PQS (see Table 4) is the one measure that uses a completely different rating method (Q-sort; Jones, 1985). The 100 items have to be sorted for each session on a continuum from 1 = least characteristic to 9 = most characteristic so that the distribution of items yield a normal distribution.

In general, all measures are based upon an external observer’s rating; the only exception is the TIQI (see Table 4) which is rated by the therapist himself. Two measures (CPPS and MULTI, see Table 4) can be additionally rated by the therapist and the patient. These two measures therefore allow simultaneous ratings from three different perspectives.

Psychometric Aspects

Objectivity

Of the 34 included measures, only 14 report a detailed coding manual which is published and/or available in the World Wide Web or from the authors and therefore fulfill the A-criterion within this category. Within these coding manuals, the definitions of the categories or items are described in detail and rater instructions are given. Only two manuals are less exact, one gives general information about the structure and use of the measure (APS, see Table 2) and one manual only gives short definitions of the categories (PIRS, see Table 2). For eight measures, the coding manual exists as an unpublished manuscript. For 13 measures, no coding instruction is reported, but a description of the scales or categories can mostly be found in the published articles.

TABLE 2 | Microanalytic measures referring to psychodynamic therapy.

| Instrument source | Basis Theoretical orientation Objective within the development | Short description | Assessment mode Raters Definition of scales | Valuation | | |
|------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|----------------------|---|
| | | | | Objectivity | Reliability Validity | |
| Analytic process scales (APS; Waldron et al., 2004a) | 18 years of bottom-up development, regular meetings over several years; PIRS (Cooper and Bond, 1992; Milbrath et al., 1999), VPPS (O'Malley et al., 1989), TVII (Koenigsberg et al., 1985, 1988), Psychotherapy Process Q-Set (Jones and Windholz, 1990) AP Examine changes occurring over time in psycho-analytic treatments and what features of patient and analyst account for these changes | 4 categories: 1. types of intervention (encourage elaboration, clarify, make an interpretation, provide support) 2. aims of intervention (patient's defense, patient's reaction to the analyst/analytic situation, patient's conflicts, romantic/sexual issues, aggressive/hostile issues, developmental focus, self-esteem issues) 3. other characteristics (confronting, feelings of the analyst, analyst's intervention amicable, analyst's intervention hostile) 4. quality of the intervention (follow the patient's immediate emotional focus, good intervention) | m (5-point Likert scale), segment of session O APS Coding Manual (80 pages) available at www.ipa.org.uk or from the author: woodywald@earthlink.net | A | B | B |
| Coding of therapist statement (Connolly et al., 1998) | / SE To classify therapist statements into general categories consistent with the techniques of SE therapy | 4 categories: 1. interpretation 2. clarification 3. question 4. other; for interpretation, additionally time frame (childhood past, adult past, present) and persons included (therapist, parents, significant other, self, no other person) are detected | d, speaking turn O No coding manual published; definition of categories in (Connolly et al., 1998) | C | A | B |
| Malan intervention typology (MIT; Malan, 1963, cited from Silberschatz et al., 1986) | / PDT To assess interpretations and their impact on therapy outcome | 2 scales: 1. interpretation (parent, other, self; transference vs. non-transference) 2. non-interpretation | d, statement O No coding manual published; description within (Malan, 1963, 1976) | C | B | C |
| Psychotherapy Interaction Coding System (PIC; McCullough, 1988, cited from Town et al., 2012b) | / PDT To examine therapist interventions and the impact of these interventions in dynamic orientated therapies | 2 scales: 1. 8 process codes: question, information, self-disclosure, clarification, directive, support, interpretation, confrontation 2. 4 content codes: defenses, anxieties, impulse/feeling, no content | d, response O Coding manual: (McCullough, 1988); description of categories: (Town et al., 2012a) | A | B | C |

(Continued)

TABLE 2 | Continued

| Instrument source | Basis Theoretical orientation Objective within the development | Short description | Assessment mode Raters Definition of scales | Valuation | | |
|----------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|-------------|----------|
| | | | | Objectivity | Reliability | Validity |
| Psychodynamic Interventions Rating Scale (PIRS; Cooper and Bond, 1992, cited from Milbrath et al., 1999) | Among others: ITS (Gaston and Ring, 1992) PDT To examine microprocesses in expressive and psychodynamic therapies regarding specific types of interpretive interventions | 2 scales: 1. interpretive interventions: defense, transference 2. supportive interventions: acknowledgments, clarification, questions, associations, reflections, work-enhancing strategies, supporting strategies, contractual arrangements | d, m for interpretative interventions (5-point Likert scale), utterance O Coding manual exists as unpub-fished manual; categories within (Eanon et al., 2001) | B | A | B |
| Therapist intervention rating system (TIRS; Piper et al., 1987) | Review of existing rating scales, (e.g., (Strupp, 1966; Luborsky et al., 1979; Marziali and Sullivan, 1980); PDT) To reliably identify interpretations (intervention defined as "interpretive" when there is reference to a "dynamic component") | 10 categories: 1. noninterventions 4 noninterpretive lower categories without reference to patient's experience: 2. formal 3. information providing 4. information requesting 5. directive 1 noninterpretive lower category (with partial reference to patient's experience): 6. nondynamic component 4 interpretive upper categories with reference to patient's internal conflict: 7. single dynamic component 8. double dynamic component 9. triple dynamic component 10. quadruple dynamic component (dynamic components: wish, fear, defensive process, dynamic expression) | d, statement O Coding manual exists as unpublished manuscript, available from the first author; description of categories with examples within (Piper et al., 1987) | A | B | C |

m, metric; d, dichotomous; O, observer; SE, supportive-expressive dynamic psychotherapy; AP, analytic psychotherapy; PDT, psychodynamic therapy; PIRS, Psychodynamic Intervention Rating Scale; VPPS, Vanderbilt Psychotherapy Process Scale; TVI, Therapist verbal Interventions Inventory. Evaluation criteria: A, no limitations; B, some restrictions; C, severe deficits (see text of the manuscript for details).

TABLE 3 | Measures referring to cognitive-behavioral orientated therapy or to a specific setting.

| Instrument source | Basis Theoretical orientation Objective within the development | Short description | Assessment mode Raters Definition of scales | Valuation | | |
|---------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|-------------|----------|
| | | | | Objectivity | Reliability | Validity |
| BEHAVIORAL ORIENTATED THERAPY—GLOBAL CODING | | | | | | |
| Cognitive Therapy Adherence-Competence Scale developed within the NIDA-CCTS (CTACS; Liese et al., 1995, cited from Barber et al., 2003) | Cognitive Therapy Scale (CTS; Young and Beck, 1980), CSPRS (Hollon et al., unpublished work), Current CT treatment manuals, format based on PACS-SE (Barber and Crits-Christoph, 1996) CT To provide a wider coverage of cognitive therapists' activities | 21 items, no subscales e.g., guided discovery, asking for evidence/alternative views, eliciting automatic thoughts, addressing key issues each item is rated separately for AD and COM | m (7-point Likert scale), session O No coding manual published, description of scales within (Barber et al., 2003) | C | A | B |
| Cognitive Therapy Scale(CTS-R; Young and Beck, 1980, cited from Blackburn et al., 2001 diverse revisions, among others by Blackburn et al., 2001) | cognitive therapy manual (Beck et al., 1979) CT to measure competence in CT (for depression) | 11 items, 2 subscales: 1. general skills subscale: agenda setting, eliciting feedback, collaboration, pacing and efficient use of time, emphatic skills, professionalism 2. specific cognitive therapy skills subscale: guided discovery, conceptualization, focus on key cognitions, application of cognitive techniques, application of behavioral techniques, use of homework | m (7-point scale), session O Coding manual available at: http://members.academyofct.org/files/documentlibrary/CTRS%281%29.pdf | A | A | B |
| BEHAVIORAL ORIENTATED THERAPY—MICROANALYTIC CODING | | | | | | |
| Coding system for the interaction in Psychotherapy (GIP; Schindler et al., 1989) | Review of literature; Verbal Response Modes (Stiles, 1978); Response Mode System (Snyder, 1945; Strupp, 1957; Hill et al., 1981; Elliott et al., 1982; Hahlweg et al., 1984) CBT To assess the patient-therapist interaction in behavior therapy | 19 items for therapist verbal behavior modes: 6 therapist dimensions: empathy (3 items), support (3 items), exploration (2 items), explanation (4 items), directivity (2 items), classification (3 items), 2 further categories: silence, remainder category | d, sentence O Coding manual existent as an unpublished manual; description of categories within (Schindler et al., 1989) | B | B | C |
| SPECIFIC SETTING | | | | | | |
| Adherence/competence scale for IDC for cocaine Dependence developed within the NIDA-CCTS (ACS-IDCCD; Barber et al., 1996) | Individual drug counseling (IDC) manual (Mercer and Woody, 1992) IDC To assess treatment AD and COM in individual drug counseling (IDC) for the treatment of cocaine-dependence | 43 items, 5 subscales: 1. monitoring drug use behavior (6 items), 2. encouraging abstinence (8 items), 3. encouraging 12-step participation (5 items) 4. relapse prevention (5 items) 5. educating the client (5 items) 1 additional section: psychodynamic/cognitive interventions (4 items) 10 additional items without categorization each item is rated separately for AD and COM | m (7-point Likert scale), session O Coding manual available from the author | A | A | B |

(Continued)

TABLE 3 | Continued

| Instrument source | Basis Theoretical orientation Objective within the development | Short description | Assessment mode | | Valuation | |
|-------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|-------------|-------------|
| | | | Raters | Definition of scales | Objectivity | Reliability |
| Sequential Code for Observing Process Change (SCOPE; Moyers and Martin, 2006) | Motivational Interviewing Skill Code (MISC; Moyers et al., 2003; Miller et al., 2008), Commitment Language Coding System (Amrhein et al., 2003) MI to evaluate the quality of motivational interviewing, to measure transition probabilities and to categorize MI-relevant Verbal behavior between therapist behavior and client behavior during MI sessions | 46 items (30 therapist, 16 client items), 8 categories: client categories: 1. neutral utterances, 2. positive commitment language (change-talk), 3. negative commitment language (sustain-talk) counselors categories: 1. questions (open vs. closed) 2. reflections (simple vs. complex), 3. MI-consistent (affirmations, emphasizing the control of the client, seeking permission prior to providing information or advice, offering support), 4. MI-inconsistent (giving advice without permission, confronting clients with information about the target behavior, directing certain courses of action, stating an opinion about client behavior or attitudes, warning clients about consequences of the target behavior) 5. other | d, utterance O MISC coding manual available at: casaa.unm.edu | A | B | B |
| Yale Adherence and Competence Scale II (YACS; Carroll et al., 2000) | Review of session videotapes and treatment manuals To rate therapist AD and COM in delivering behavioral treatments for substance use disorders | 40 items, 6 subscales: 1. assessment (5 items), 2. general support (5 items), 3. goals of treatment (5 items), 4. clinical management (CM, 10 items), 5. twelve step facilitation (TSF, 9 items), 6. cognitive-behavioral treatment (CBT, 6 items) each item is rated separately for AD and COM | m (5-point Likert scale), session O Coding manual available at: http://www.mirecc.va.gov/vish1/docs/products/Yale_Adherence_and_Competence_Scale_II_Guidelines.pdf | A | A | A |

m, metric; d, dichotomous; O, observer; CT, cognitive therapy; CM, clinical management; AD, adherence; COM, competence; CBT, cognitive behavioral therapy; TSF, twelve step facilitation; IDC, individual drug counseling; SD, standard deviation; SE, supportive expressive therapy; NIDA-CCCTS, National Institute on Drug Abuse-Collaborative Cocaine Treatment Study; CSPRS, Collaboratory Study Psychotherapy Rating Scale; PACS-SE, Penn Adherence-Competence Scale for Supportive-Expressive Therapy; MI, motivational interviewing. Evaluation criteria: A, no limitations; B, some restrictions; C, severe deficits (see text of the manuscript for details).

TABLE 4 | Pantheoretical global measures.

| Instrument source | Basis Theoretical orientation Objective within the development | Short description | Assessment mode Raters Definition of scales | Valuation | |
|--------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|----------------------|
| | | | | Objectivity | Reliability Validity |
| Comparative Psychotherapy Process Scale (CPPS; Hilsenroth et al., 2005) | Two reviews of the empirical comparative psychotherapy process literature (Blagys and Hilsenroth, 2000, 2002) PI, CBT To classify and compare PI and CBT treatments and to assess therapist activity, process and techniques used in a therapy session | 20 items, 2 subscales: 1. psychodynamic-interpersonal subscale (10 items), 2. cognitive-behavioral subscale (10 items) | m (7-point Likert scale), session P,T,O Coding manual available at http://supp.apa.org/psycarticles/supplemental/pst_42_3_340/pst_42_3_340_supp.html | A | A A |
| Collaborative Study Psychotherapy Rating Scale (Form 6) (CSPRS; Hollon et al., 1984, cited from Hill et al., 1992) | Consultation of trainers of treatment modalities, several treatment manuals, based on MTRS (DeRubeis et al., 1982) CBT, IPT, CM To distinguish sessions of cognitive therapy, interpersonal psychotherapy, and clinical management pharmacotherapy | 96 items, 8 scales: 3 modality-specific scales: 1. CBT scale (28 items, 6 subscales), 2. IPT scale (28 items, 7 subscales), 3. CM scale (20 items, 5 subscales) 3 tangential modality scales (8 items, 2 CBT, 1 IPT subscale) 2 non-modality-specific scales (12 items, 2 subscales: facilitative condition (FC, 8 items), explicit directiveness (ED, 4 items) | m (7-point Likert scale), session O Coding manual (Evans et al., 1984) | A | A A B |
| Sheffield Psychotherapy Rating Scale (SPRS; Shapiro and Startup, 1990, cited from Startup and Shapiro, 1993) | 41 items of CSPRS, (3 items slightly changed), 16 new items ET, CBT, IPT Designed to rate adherence of therapists to exploratory therapy (psychodynamic/experiential therapy with an interpersonal focus) | 59 items, 3 scales: 1. exploratory therapy scale (ET, 19 items (4 items from the IPT scale of CSPRS, 14 new items) 6 subscales) 2. prescriptive scale (P, 32 items; CBT + 2 of tangential modality scale of CSPRS + 2 new items) 3. FC scale of CSPRS (8 items) | m (7-point Likert scale), session O Raters' manual (Shapiro and Startup, 1990), description of items within (Startup and Shapiro, 1993) | A | A A B |
| Multitheoretical List of Therapeutic Interventions (MULTI; McCarthy and Barber, 2009) | Treatment manuals, therapy books, adherence measures, theoretical and review articles, experts To assess therapeutic interventions from different psychotherapy orientations and from the perspective of patients, therapists, and observers | 60 items, 8 subscales: 1. behavioral (BT, 15 items), 2. cognitive (CT, 16 items), 3. dialectical-behavioral (DBT, 8 items), 4. interpersonal (IPT, 7 items), 5. person centered (PC, 7 items), 6. psychodynamic (PD, 12 items), 7. process-experiential (PE, 9 items), 8. common factors (CF, 7 items) (several items are unspecific and belong to multiple subscales) | m (5-point rating scale), session P,T,O No coding manual published | C | A A B |

(Continued)

TABLE 4 | Continued

| Instrument source | Basis Theoretical orientation Objective within the development | Short description | Assessment mode | | Valuation | |
|------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------|-------------|-------------|
| | | | Raters | Definition of scales | Objectivity | Reliability |
| Psychotherapy process O set (POS; Jones, 1985, cited from Jones et al., 1987) | Search of extant process measures, discussions with research-oriented clinicians, bottom-up development To provide a standard language and rating procedure for classification of the therapy process and to systematically characterize a wide range of patient-therapist interactions | 100 items, no subscales: 3 types of items: 1. patient attitude and behavior or experience (40 items), 2. therapist's actions and attitudes (41 items), 3. nature of the interaction in the dyad or the climate or atmosphere (19 items) | Q method: items are sorted on a continuum from 1=least to 9=most characteristic (5=neutral/irrelevant), session O Coding manual exists as unpub-lished manual (Jones, 1985) | B | B | B |
| Therapist Interventions and Qualities Inventory – therapist form (TIQI-T) (TIQI-P; patient form Bogwald, 2001) | Previous research of literature, feedback of experienced clinicians, Barrett-Lennard Relationship Inventory (1962) To assess specific therapist interventions and unspe-cific interpersonal skills and qualities that have been emphasized in other instruments e.g., Barrett-Lennard Relationship Inventory | 36 items, no subscales: e.g.: encouraged the patient to explore uncomfortable emotions, made interpretations, encouraged patient to talk about what others might feel toward patient, questioned patient about his/her feeling toward therapist | m (5-point Likert scale), session T No coding manual published | C | B | C |
| Vanderbilt Psychotherapy Process Scale Developed within Vanderbilt Psychotherapy Research Project (PPS; O'Malley et al., 1983; Strauss et al., 1992) | Based on <i>Therapy Session Report</i> (O'flinsky and Howard, 1967, 1975) and further development; (Gomes-Schwartz and Schwartz, 1978) To assess relevant aspects within the therapeutic process Three domains: exploratory processes, therapist-offered relationship, and client involvement | 64 items, 8 subscales: 1. client participation (8 items), 2. client hostility (6 items), 3. client dependency (6 items), 4. therapist warmth and friendliness (9 items), 5. negative therapist attitude (6 items), 6. client exploration (7 items), 7. therapist exploration (13 items), 8. client psychic distress (9 items) | m, 15-minute segment of session or entire session O Coding manual existing as an unpublished manuscript | B | B | B |

m, metric; P, patient; T, therapist; O, observer; CBT, cognitive-behavioral therapy; PI, psychodynamic-interpersonal therapy; IPT, interpersonal therapy; CM, clinical management; ET, exploratory therapy; FC, Facilitative conditions. Evaluation criteria: A, no limitations; B, some restrictions; C, severe deficits (see text of the manuscript for details).

TABLE 5 | Pantheoretical microanalytic measures.

| Instrument source | Basis Theoretical orientation Objective within the development | Short description | Assessment mode Raters Definition of scales | Valuation | | |
|-------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|-------------|----------|
| | | | | Objectivity | Reliability | Validity |
| Coding System of Therapeutic Focus (CSTF; Goldfried et al., 1989, cited from Kerr et al., 1992) | Consultation of researchers /practitioners from CBT and IPT; preliminary scoring of therapeutic vignettes from other psychotherapy researchers and published transcripts appearing in the literature <i>CBT, PI</i> To examine the content of the therapist's in-session focus and to facilitate comparative process analyses across therapeutic orientations | 40 items, 6 categories: 1. components of functioning (10 items), 2. general interventions (7 items), 3. intrapersonal links (3 items), 4. interpersonal links (5 items), 5. persons involved (9 items), 6. time frame (6 items) | d, utterance O Original article is an unpublished manuscript; for description of scales in detail, see e.g., (Goldsamt et al., 1992); no coding manual published | C | A | B |
| Helping Skills System (HSS; Hill and O'Brien, 1999, cited from Hess et al., 2006) | Substantially revised version of the HCVRCS (Hill Counselor Verbal Response Category System, Hill, 1978) To classify counselor's responses without regard to the quality of the message | 12 categories; 3 subscales: 1. exploration skills (approval/reassurance, closed question, open question, restatement, reflection of feelings), 2. insight skills (challenge, interpretation, self-disclosure, immediacy), 3. action skills (information, direct guidance). residual category | d, grammatical sentence O Coding manual available at http://forms.apa.org/books/supp/hill3/pdf/student/webformE.pdf | A | B | B |
| Hill Counselor Verbal Response Category System-Revised (HCVRCS-R; Friedlander, 1982) | Revised version of HCVRCS (Hill, 1978) Conceptually and methodologically more rigorous than the original version | 9 categories: encouragement/approval/reassurance, reflection/restatement, self-disclosure, confrontation, interpretation, providing information, information seeking, direct guidance/ advice, unclassifiable | d, utterance O No coding manual exists, description within (Friedlander, 1982) | C | B | C |
| Inventory of Therapeutic Strategies (ITS; Gaston and Ring, 1992) | Review of literature, review of therapy sessions (of cognitive, dynamic, and behavioral therapy) To assess the major intentions in therapist's interventions | 19 items, 3 subscales: 1. explorative strategies (12 items, 3 content subcategories (defenses, emotions, cognitions), 4 objective subcategories (therapist, others, self, non-interpersonal situations) 2. supportive interventions (3 items), 3. work-enhancing strategies (4 items) | microanalytic: d, statement; global: m (5-point Likert scale), session O Coding manual exists as unpublished ms. (Gaston, 1988) | B | A | A |
| Helper Behavior Rating System (Shapiro et al., 1980), c.f.r. (Barkham and Shapiro, 1986) | Elliott's (1979) Helper Behavior Rating System (Shapiro et al., 1984), review of existing rating systems To develop a system of response mode analysis for clinical research | 12 categories: interpretation, exploration, reflection, general advisement, process advisement, reassurance, disagreement, open question, closed question, general information, self-disclosure, other | d, thought unit O Coding manual (Shapiro et al., 1980), available from the senior author | A | B | C |

(Continued)

TABLE 5 | Continued

| Instrument source | Basis Theoretical orientation Objective within the development | Short description | Assessment mode Raters Definition of scales | Valuation | | |
|-----------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------|-------------|-------------|----------|
| | | | | Objectivity | Reliability | Validity |
| Response modes coding system (Connolly Gibbons et al., 2002) | Based on 6 published coding systems described within (Elliott et al., 1987) IPT, CT To compare and contrast the therapist response modes used in manual-guided interpersonal and cognitive therapy sessions | 3 categories, 8 subcategories: 1. statement categories: learning statements, clarification, restatements, questions, informational/directional statements, self-disclosures, role play, other statements 2. time frames: childhood, adult past, present, in session, future, unspecified 3. person code: father, mother, parents, sibling, immediate family, extended family, therapist, other real person, unspecified, significant other, child, patient | d, speaking turn O Manual existing as unpublished manuscript, description of categories within (Connolly Gibbons et al., 2002) | C | B | C |
| System for assessing therapist communications (SATC; Brunink and Schroeder, 1979) | System Of Analysis (SOA; Fiedler, 1950, 1951; Strupp, 1957; Carikhuff, 1969) To compare psychoanalytically orientated, behavior, and Gestalt therapists | 6 dimensions: type of therapeutic activity (8 categories: facilitating communications, exploratory operations, clarification, interpret-tive operations, structuring, direct guidance, activity not clearly relevant to task of therapy, and unclassifiable), temporal focus (present, past), interview content focus (4 categories: client, therapist-client relationship, therapist self-disclosures, content irrelevant to therapy, degree of initiative, communication (5 categories: relevant task-oriented communications, accurate, nonadditive comm., additive comm., inaccurate or subtractive comm., comm. irrelevant to therapy, therapeutic climate (4 categories: minimally supportive or emotionally neutral, supportive or giving, nonsupport-tive or withholding) | d (m for "degree of initiative," 4 levels), utterance O No coding manual published; description of dimensions within (Brunink and Schroeder, 1979) | C | B | B |
| Therapist Behavior Code-Revised (TBC-R; Bischoff and Tracey, 1995) | (TBC; Forigatch and Chamberlain, 1982) Focus on categorization of therapist's behavior into directive and nondirective units | 8 categories: support (10 items), teach (5 items), structure (3 items), question and information seek (4 items), confront and challenge (8 items), interpret and reframe (6 items), talk (1 item), facilitate (1 item); additional subcodes for question and interpret: directive vs. non-directive | d, speaking turn O Coding manual existent as an unpublished manuscript | B | B | C |

(Continued)

TABLE 5 | Continued

| Instrument source | Basis Theoretical orientation Objective within the development | Short description | Assessment mode Raters Definition of scales | Valuation | | |
|-------------------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|---------------------------------------------------|-------------|-------------|----------|
| | | | | Objectivity | Reliability | Validity |
| Verbal Response Mode (VRM; Stiles et al., 1989) | First author's previous work CBT, PI To classify speech acts which concern what people do when they say something, without respect to speech content or to interactants' affective states (version for client and for therapist available) | 8 categories: disclosure (D), edification (E), question (Q), acknowledgment (K), advisement (general/process) (A), confirmation (C), interpretation (I), reflection (exploratory/simple) (R) Each category is rated separately regarding interpersonal intent and grammatical form | d, utterance O Coding manual (Stiles, 1992) | A | A | C |

m, metric; d, dichotomous; O, observer; CT, cognitive therapy; CBT, cognitive-behavioral therapy; PI, psychodynamic-interpersonal therapy; IPT, interpersonal therapy. Evaluation criteria: A, no limitations; B, some restrictions; C, severe deficits (see text of the manuscript for details).

Reliability

There are large differences regarding the reliability of the included measures (see **Tables 1–5**, column 9 and 10; see Supplementary Material for procedure and results of reliability calculation). First, the procedures of reliability calculation differ widely. The number of sessions and patients on which the values are based ranges from three sessions with three patients (total of 9 sessions; Silberschatz et al., 1986) to eight sessions with 33 clients (total of 264 sessions; Stiles et al., 1989). Some studies report reliability scores based on consensus ratings (e.g., Butler et al., 1995) or ratings achieved by regular meetings to prevent rater drift (e.g., Gaston and Ring, 1992; Connolly Gibbons et al., 2002). Some reliability values are based on ratings of students, others on ratings of the measures' authors. In most cases, raters have been trained extensively before using the measure (e.g., Goldfried et al., 1997), but sometimes not (e.g., McCarthy and Barber, 2009). Some studies describe the rating procedure incompletely so that information about the number of rated sessions is missing.

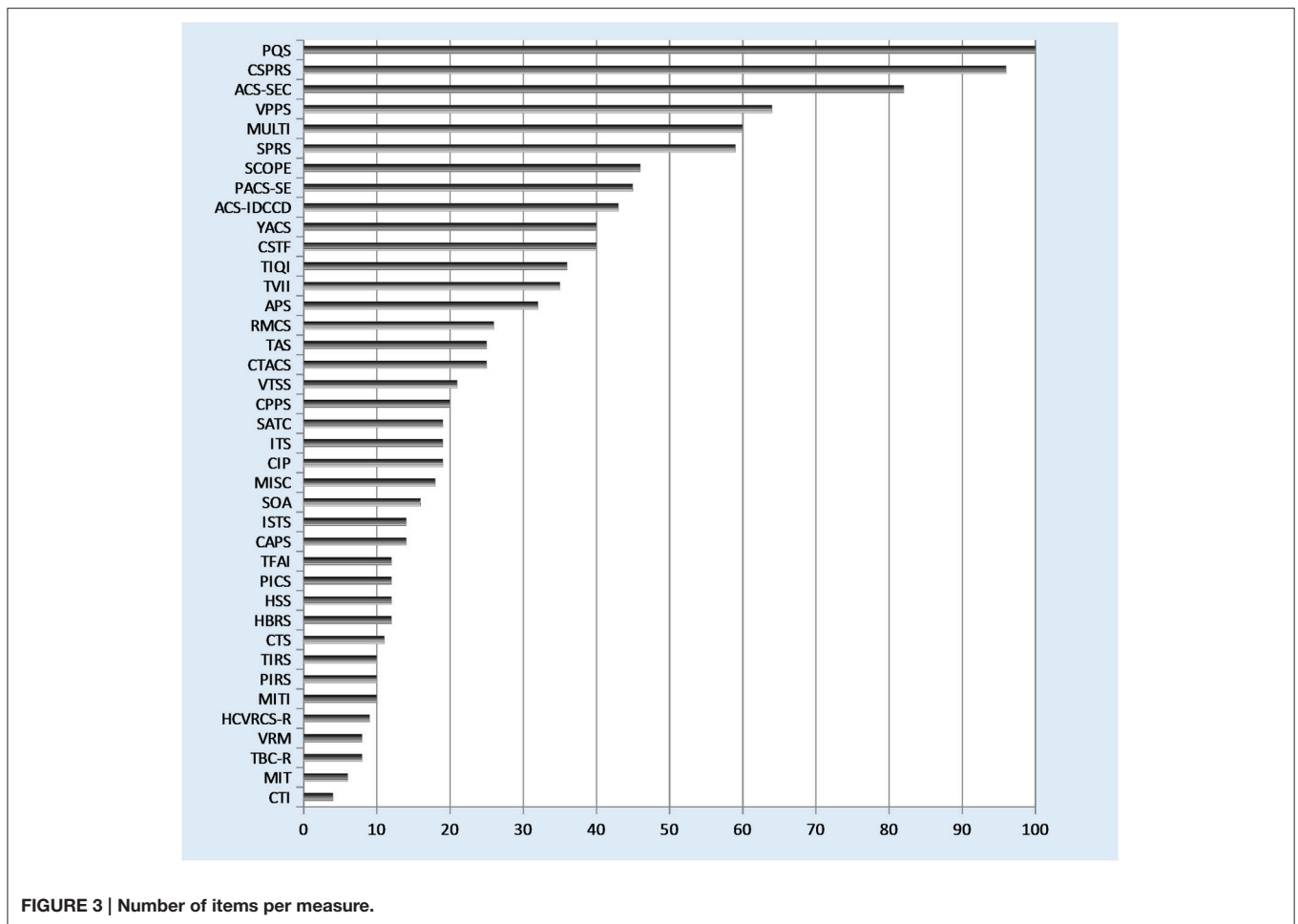
Secondly, the reliability gold standard by calculating ICC's is fulfilled by $n = 16$ of the included measures. Within these, ICC (2,2) and ICC (2,1) are mostly used. For two measures, it remains unclear which form of ICC was calculated (VTSS; see **Table 1**, Butler et al., 1995; CTS-R; see **Table 3**, Blackburn et al., 2001). Overall, ICC values range from acceptable to high levels. Scales, categories or items with ICC values below the acceptable level are found in measures with discrete categories, so none of the measures had to be excluded because of lacking reliability, although ICC values are presented.

All other measures ($n = 18$) report Finn's r correlation coefficients, Pearson product-moment correlation coefficients or kappa values. Similar to ICC values, it is often not clearly specified which kind of kappa has been applied (e.g., "Light's kappa," Light, 1971; "Fleiss kappa," Fleiss, 1971), whether it is the median or the average kappa). One measure failed to report reliability values for all subscales. Since available values were good, the measure was not excluded from the review.

Validity

Six of the included measures report no information about the validity of the measure and three further measures report insufficient information. For three measures, studies are available which report a broad examination and evidence for validity for the respective measure: CPPS (see **Table 4**; Hilsenroth et al., 2005), ITS (see **Table 5**; Gaston and Ring, 1992), and YACS (see **Table 3**; Carroll et al., 2000). All other measures report validity on a minimal level, especially with regard to criterion validity (which means that scales reflect and distinguish between different therapeutic orientations) and convergent validity (which means that the scale has been compared with another measure).

All in all, only two measures fulfill the A-criterion for all three categories (CPPS, see **Table 4**, and YACS, see **Table 3**; see Supplementary Material for procedure and results of validity calculation).



DISCUSSION

The objective of the present study was to review available measures designed to assess directly observable (i.e., formal, temporal, and thematic) dimensions of verbal therapeutic techniques in a reliable and valid manner. All in all, 34 measures were identified which show a great heterogeneity regarding purposes, theoretical foundations and assessment procedures. Whereas there is much more consensus related to outcome measures in psychotherapy (e.g., Strupp et al., 1997). In general, it appears that most research groups rather created their own new measure than using existing ones in order to address their research questions. Explanations of this approach include that available measures are too specific in focus (McCarthy and Barber, 2009), do not cover the necessary constructs under investigation (Milbrath et al., 1999), require too much effort, or show insufficient psychometric properties (Hilsenroth et al., 2005). This practice, however, generated a number of problems in the assessment of verbal techniques.

Issues Regarding the Systematization of Techniques

The reviewed measures refer to a specific theoretical orientation or to different theoretical orientations. They have been developed to address different research questions such as assessing

adherence and competence of the therapist regarding specific therapeutic orientations, distinguishing between therapeutic orientations, investigating microprocesses in psychotherapy, or analyzing the relationship between techniques and therapy outcome.

Some of them assess all possible verbal techniques (again within a therapeutic orientation or in general), whereas others focus on selected theoretically important constructs such as “interpretation.” “Interpretation” is the category or item that is assessed by the majority of measures. The focus is partly limited to only one or two aspects of verbal techniques (e.g., differentiation between “interpretation” vs. “non-interpretation” or “supportive techniques” vs. “interpretative techniques”). There are heterogeneous levels on which techniques are differentiated. The number of categories ranges from four categories without further differentiation to 100 items without further categorization.

A potential problem is that the hierarchical structures of the measures are different, although the same or similar techniques are assessed. For example, the ISTS (see **Table 1**) contains the scale “interpretative techniques” (level 1) which includes further subordinated items like the “impression of the therapist” (level 2). The HSS (see **Table 5**), on the other hand, subsumes items as “interpretation” (level 2) within the scale “insight skills” (level 1). Thus, it is difficult to specify the relationship of scales and items

in both measures (notably the intervention “interpretation”) as well as results derived from using them.

The definitions of verbal techniques are also very different across measures. In some cases, the same categories are defined in different ways. For example, the category “interpretation” was defined as a therapeutic utterance which goes beyond the perceptions of the patient (Hill, 1978), an utterance which refers to one or several dynamic components (e.g., wishes, fears, defense mechanisms; Piper et al., 1998) or a theory-derived response of the analyst, rated according to the degree to which it transforms meaning by bringing aspects outside of awareness into full awareness (Waldron et al., 2004b). In other cases, intervention categories show overlaps and can hardly be distinguished from one another. This situation is certainly due to and aggravated by the fact that—in the theoretical literature—specific intervention techniques are described and defined inconsistently, abstractly and in part ambiguously, which limits their operationalizability and investigation in evidence-based therapy research (Fonagy, 2000; Paniagua, 2003; Gumz et al., 2014).

Another problem derives from the different facets of the assessment of verbal techniques. One problem in this respect is that many items or subscales of available measures simultaneously assess different features of therapeutic techniques (e.g., formal and qualitative features, thematic content, and therapist’s intentions), prohibiting the possibility to analyze these features separately (e.g., VTSS, see **Table 1**: “Therapist encourages the patient’s expression and/or exploration of feelings in relation to a significant other (including therapist.”)

From our point of view it is helpful to clearly specify on which level techniques are to be described. There is a difference between settings in which more global descriptions are needed (e.g., role play technique, sculpture work, hypnosis or transference regarding a more comprehensive topic) and settings in which a focus on the microanalytic level of single verbal interventions is more important (e.g., the verbalization of emotional reactions or a specific transference interpretation; Gumz et al., 2014) and where there is greater overlap between different therapeutic schools. Furthermore, it would be beneficial to separate the basic features of verbal techniques (Form, Temporal focus, Thematic content, Latent characteristics). This relates specifically to the latent characteristics. This would provide the opportunity to evaluate different aspects of therapeutic techniques, i.e., to differentiate between what the therapist said and what his utterance implied. Even if this strict systematic may be difficult to realize in research it may be useful to compare results and conduct meta-analyses.

Issues Regarding Assessment Mode and Rating Perspective

Half of the measures reviewed ($n = 17$) use global assessment (i.e., rating of complete sessions) while the other 17 measures use a microanalytic assessment referring to different units like utterances, sentences or speaking turns. The advantage of global assessment is that they allow for individual weighting of information (Bøgwald et al., 1999) and that they are time-saving, thereby allowing the analysis of larger sets of data. On the other hand, global methods deliver rather crude assessments of

complex processes and are, moreover, prone to cognitive biases, e.g., anchoring effects and availability heuristics (Tversky and Kahneman, 1974; see also Heaton et al., 2010). Microanalytic assessments, on the other hand, are more labor-intensive which limits the amount of data to be analyzed. However, microanalytic methods allow for the analysis of the effects of specific intervention techniques on a micro-level which is to date hardly known and would help to clarify the processes involved in psychotherapeutic change. In summary, both assessment modes have advantages as well as disadvantages that need to be considered before conducting a study. A general desideratum in this respect will be to define what “a verbal intervention” actually is, i.e., which unit (single sentence vs. speaking turn vs. therapy session) needs to be assessed.

Another issue refers to the rating perspective. Except for three measures (TIQI, CPPS, MULTI, see **Table 4**), all included measures have to be used by an external observer. This perspective may be sufficient to answer many research questions regarding verbal techniques. Some topics, however, cannot be covered by analyzing external ratings, e.g., differences in the perception of verbal techniques by therapist vs. patient. Accordingly, it would be helpful to conceive measures which allow the assessment of different perspectives on the therapy process to increase the range of analytic possibilities.

Issues Regarding Psychometric Aspects

A final set of problems relates to the psychometric properties of the measures reviewed. Objectivity is the least problematic aspect in this respect: For most measures ($n = 22$) the authors reported the availability of explicit coding manuals. Some of these manuals were published ($n = 14$) which makes it possible to reconstruct the rating process in detail (A-criterion). In unpublished manuals ($n = 7$), it is more difficult to draw firm conclusions about the objectivity of the rating process. However, it can be assumed that the ratings within published studies are based on manuscripts describing the use of the measures so that objectivity is given on a minimal level (B-criterion). For the remaining measures ($n = 13$), no manuals were reported so that the objectivity of the rating process is not guaranteed.

The reliability of the measures reviewed is much more difficult to evaluate, since different authors used a wide range of different methods of reliability assessment. For $n = 16$ measures ICC values were reported and classified as sufficiently reliable (A-criterion). The remaining measures ($n = 18$) show a satisfying reliability reporting kappa, Fleiss’ r or product-moment correlations. It was observed that reliability values in some cases should be treated with caution, because the rating procedures were described incompletely, the exact method of reliability assessment was not specified (e.g., “kappa” instead of “Cohen’s kappa” or “Fleiss” “kappa”), or the occurrence of interrater discussions which may lead to a falsifying increase of reliability over time. The evaluation of the validity of results proved to be the most problematic aspect in the assessment of verbal techniques. Only three measures (CPPS, see **Table 4**; ITS, see **Table 5**; YACS, see **Table 3**) revealed good values for different fields of validity while for all other measures there is only minimal or hardly any information related to validity. The

question remains, why the important validation criterion is not considered sufficiently within the scale development process.

All in all, only two measures fulfilled the A-criterion in matters of objectivity, reliability and validity, i.e., the CPPS (see **Table 4**, which allows to discriminate between two or more therapeutic orientations) and the YACS (see **Table 3**, which aims to assess the adherence and competence of the therapist in treatments for substance use disorders). From a methodological perspective, these are the measures which can be recommended for clinical use regarding the respective research questions without reservation. All other measures show more or less severe psychometric deficits regarding psychometric aspects. From a practical perspective, some of these measures might nevertheless be helpful to address specific research questions, particularly to analyze the relationship between therapeutic techniques and therapy outcome or to investigate microprocesses in psychotherapy. However, in general it needs to be stated that further research has to operate with more refined methods ensuring the objective, reliable and valid rating of verbal therapeutic techniques.

Limitations

It could be possible that the used keywords did not identify all instruments that assess verbal techniques. Another limitation is that only papers in English and German language were considered which excludes measures in other languages. Investigating the effect of techniques on therapeutic outcome was outside the scope of this review, and would be a valuable contribution for future research.

Are the Proposed Standards Reasonable or a Pie in the Sky?

Our systematic review revealed a number of problems regarding the measures available for the assessment of verbal techniques. Due to these problems of assessing verbal techniques, it is almost impossible to compare the scales and categories of different measures as well as results derived from studies, even if they examine the same outcome variable. Accordingly, no definite conclusions regarding their efficacy can be drawn.

Where does this leave us? In this review, we have certainly used very high standards for measures of psychotherapy research. It must be discussed if these standards are really reasonable or if they are “pie in the sky,” since they are difficult to accomplish and may be too rigorous. Certainly, these issues might be painfully obvious to many investigators. On the other hand, it is important and useful to reiterate these criteria to enhance progress in research and to reach standards common in psychotherapy outcome research (e.g., Strupp et al., 1997). Some of the authors of the studies we rated as deficient, undoubtedly might have provided interesting and reasonable justifications as to why they did not attain what we refer to as the “gold” standard. Moreover, we have also listed some problems which cannot be avoided or need to be tolerated. Nevertheless, a solid theoretical as well as a methodological framework for measuring verbal techniques using clear terminology and accurate conception is of paramount importance to advance this field of research.

Language is the cornerstone of most psychotherapy sessions. Knowing which specific verbal techniques “good” therapists use in their practical work has a high practical relevance for psychotherapist training and clinical routine. Many questions remain in the domain of studying verbal techniques. We need more sophisticated theoretical knowledge regarding the question which techniques really matter and which should be classified as “clinical lore” (Barber, 2015, p. 325). Moreover it is important to know what the specific features of a particular technique actually are. And it will be important to achieve a consensus concerning the definition and operationalization of verbal techniques. Regarding these topics more discussions between clinicians and researchers are necessary. An optimal solution would be to define techniques without reference to therapeutic methods and to find clear unambiguous designations for specific techniques so that the same thing is not called by different names. Authors of measures need to clarify the specific features of verbal techniques that are to be assessed. Although the four basic features which we suggested (*Form, Temporal focus, Thematic content, Latent characteristics*) are not always easily distinguished, it may be worth the effort, as this would finally help to disentangle the complex mechanisms involved in the efficacy of verbal techniques.

What is the importance of “specific” techniques which are considered to belong to a particular psychotherapy method compared to other techniques which are used without reference to a particular method? In order to answer this question the creation and use of multitheoretical scales is of great theoretical and practical utility (Barber, 2015, p. 232). The question can be broadened: To what extent is the outcome due to unintended or even non-theoretically relevant techniques (Barber, 2015). Answering this latter question it would be useful to describe what therapists do without any reference to a specific theory. For this purpose, bottom-up, qualitative approaches could be helpful to develop scales in order to be able to gather theoretically unbiased, non-predetermined and comprehensive information.

Another interesting research question would be how the latent features of verbal techniques complement and interact with the directly observable features in explaining good outcome, or how the use of different therapeutic techniques changes during the course of therapy.

A significant proportion of the outcome variance is explained by a therapist effect (Wampold, 2001). Our understanding of the variables responsible for these effects of the therapist appears to be lacking (Castonguay, 2011). It may be particularly interesting to study the possible interaction between personal characteristics and therapeutic techniques. Apart from focusing on who the therapist is, what he/she should do in therapy to facilitate change is likely to be of great interest to clinicians (Castonguay, 2011).

A final important question would be how the techniques interact with alliance or other common factors to bring about patient improvement. Castonguay (2011) named the investigation of the interaction between participants, relationship, and technique variables for different clinical disorders as one of the two most important directions of future research.

Answering these clinically relevant questions can be facilitated by establishing a theoretically clear system and accurate conceptions as well as a methodically solid basis.

CONCLUSIONS

To date, the results regarding verbal interventions gathered with available instruments can hardly be compared. This has mainly three reasons: 1. Different objectives for the use of the instruments (e.g., measuring competence and adherence vs. differentiation between therapy methods vs. analyzing specific theoretical constructs), 2. Insufficient systematization (e.g., analysis of different or different numbers of interventions, partly on different hierarchical levels), and 3. Inconsistencies in the definition of categories (definitions with the same label often do not match, or vice versa, identical aspects have differing designations).

In conclusion, the overview over the status quo of assessment methods allows for the formulation of important desiderata of future research. Firstly, it will be important to achieve a consensus concerning the definition and operationalization of verbal techniques.

Secondly, we suggest that authors of measures need to clarify the specific features of verbal techniques that are to be assessed. In our review, we applied a typology comprising four features, i.e., the *Form*, the *Temporal focus*, the *Thematic content*, and *Latent characteristics*. Although these four features are not always easily distinguished, it may be worth the effort, as this would result in a more focused approach toward verbal techniques which would finally help to disentangle the complex mechanisms involved in the efficacy of verbal techniques. It would reduce the chaos and complexities of definitions and make it easier to find an orientation among concepts and empirical findings.

Thirdly, it will be important to put more emphasis on the psychometric properties of measures. Our review revealed that only two of the available measures fulfill the highest standard of the three central validation criteria. In general, authors of future measures need to provide explicit coding manuals ensuring the objectivity of the rating process and demonstrate the reliability

and validity of the assessment. Finally, our review revealed that to date there is no measure which was primarily developed in the course of a qualitative bottom-up approach, i.e., starting from empirical data of therapeutic utterances in therapy sessions which are then explicitly described and classified. For future research, we recommend that this approach should be granted more attention, as it allows the gathering of comprehensive information (Schreier, 2012).

Given that future research meets these basic desiderata, it should be possible to establish a firmer ground for the assessment of verbal therapeutic techniques. This would, as a consequence, advance the state of knowledge about the question which techniques can help whom with specific sets of symptoms in certain situations, and finally help to specify empirically based guidelines for psychotherapeutic practice.

AUTHOR CONTRIBUTIONS

AG conceived and coordinated the study. BT and HW participated in the design of the study. All authors (AG, BT, BS, CM, HW) substantially contributed to the acquisition, analysis, or interpretation of data for the work. AG, HW, and CM drafted the manuscript. All authors revised it critically for important intellectual content. All authors gave their final approval of the version to be published. All authors gave their agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

FUNDING

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://journal.frontiersin.org/article/10.3389/fpsyg.2015.01705>

REFERENCES

- Ablon, J., and Jones, E. E. (1999). Psychotherapy process in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *J. Consult. Clin. Psychol.* 67, 64–75. doi: 10.1037/0022-006X.67.1.64
- Adler, L. M., and Enelow, A. J. (1966). An instrument to measure skill in diagnostic interviewing: a teaching and evaluation tool. *J. Med. Educ.* 41, 281–288. doi: 10.1097/00001888-196603000-00013
- Amrhein, P. C., Miller, W. R., Yahne, C. E., Palmer, M., and Fulcher, L. (2003). Client commitment language during motivational interviewing predicts drug use outcomes. *J. Consult. Clin. Psychol.* 71, 862–878. doi: 10.1037/0022-006X.71.5.862
- Banon, E., Evan-Grenier, M., and Bond, M. (2001). Early transference interventions with male patients in psychotherapy. *J. Psychother. Pract. Res.* 10, 79–92.
- Barber, J. P. (2015). “Core conflicts in psychotherapy research,” in *Visions in Psychotherapy Research and Practice: Reflections from the Presidents of the Society for Psychotherapy Research*, eds B. M. Strauss, J. B. Barber, and L. G. Castonguay (New York, NY: Routledge), 316–328.
- Barber, J. P., and Crits-Christoph, P. (1996). Development of a therapist adherence/competence rating scale for supportive-expressive dynamic psychotherapy: a preliminary report. *Psychothe. Res.* 6, 81–94. doi: 10.1080/10503309612331331608
- Barber, J. P., Foltz, C., Crits-Christoph, P., and Chittams, J. (2004). Therapists’ adherence and competence and treatment discrimination in the NIDA Collaborative Cocaine Treatment Study. *J. Clin. Psychol.* 60, 29–41. doi: 10.1002/jclp.10186
- Barber, J. P., Krakauer, I., Calvo, N., Badgio, P. C., and Faude, J. (1997). Measuring adherence and competence of dynamic therapists in the treatment of cocaine dependence. *J. Psychother. Pract. Res.* 6, 12–24.
- Barber, J. P., Liese, B. S., and Abrams, M. J. (2003). Development of the cognitive therapy adherence and competence scale. *Psychother. Res.* 13, 205–221. doi: 10.1093/ptr/kpg019
- Barber, J. P., Muran, J. C., McCarthy, K. S., and Keefe, R. J. (2013). “Research on psychodynamic therapies,” in *Bergin and Garfield’s Handbook of Psychotherapy*

- and Behavior Change, 6th Edn., ed M. J. Lambert (John Wiley & Sons, Inc), 443–494.
- Barber, J. P., Mercer, D., Krakauer, I., and Calvo, N. (1996). Development of an adherence/competence rating scale for individual drug counseling. *Drug Alcohol Depend.* 43, 125–132. doi: 10.1016/S0376-8716(96)01305-1
- Barkham, M., and Shapiro, D. (1986). Counselor verbal response modes and experienced empathy. *J. Couns. Psychol.* 33, 3–10. doi: 10.1037/0022-0167.33.1.3
- Barlow, D. H. (2010). Negative effects from psychological treatments: a perspective. *Am. Psychol.* 65, 13–20. doi: 10.1037/a0015643
- Barrett-Lennard, G. T. (1962). Dimensions of therapist response as causal factors in therapeutic change. *Psychol. Monogr.* 76, 1–36. doi: 10.1037/h0093918
- Bartko, J. J., and Carpenter, W. T. (1975). On the methods and theory of reliability. *J. Nerv. Ment. Dis.* 163, 307–317. doi: 10.1097/00005053-197611000-00003
- Beck, A. T., Rush, A. J., Shaw, B. F., and Emery, G. (1979). *Cognitive Therapy of Depression: A Treatment Manual*. New York, NY: Guilford Press.
- Bergin, A. E., and Strupp, H. H. (2009). *Changing Frontiers in the Science of Psychotherapy*. Chicago, IL: Aldine-Atherton.
- Beutler, L. E. (2009). Making science matter in clinical practice: redefining psychotherapy. *Clin. Psychol.* 16, 301–317. doi: 10.1111/j.1468-2850.2009.01168.x
- Bischoff, M. M., and Tracey, T. J. G. (1995). Client resistance as predicted by therapist behavior: a study of sequential dependence. *J. Consult. Psychol.* 42, 487–495. doi: 10.1037/0022-0167.42.4.487
- Blackburn, I.-M., James, I. A., Milne, D. L., Baker, C., Standart, S., Garland, A., et al. (2001). The Revised Cognitive Therapy Scale (CTS-R): psychometric properties. *Behav. Cogn. Psychother.* 29, 431–446. doi: 10.1017/s1352465801004040
- Blagys, M. D., and Hilsenroth, M. J. (2000). Distinctive features of short-term psychodynamic-interpersonal psychotherapy: a review of the comparative psychotherapy process literature. *Clin. Psychol.* 7, 167–188. doi: 10.1093/clipsy.7.2.167
- Blagys, M. D., and Hilsenroth, M. J. (2002). Distinctive activities of cognitive-behavioral therapy - A review of the comparative psychotherapy process literature. *Clin. Psychol. Rev.* 22, 671–706. doi: 10.1016/S0272-7358(01)00117-9
- Bøgwald, K.-P. (2001). Do patients and their therapists agree on the content of treatment? *J. Nerv. Ment. Dis.* 189, 830–837. doi: 10.1097/00005053-200112000-00004
- Bøgwald, K.-P., Høglend, P., and Sørbye, O. (1999). Measurement of transference interpretations. *J. Psychother. Pract. Res.* 8, 264–273.
- Bordin, E. S. (1966). “Free association: an experimental analogue of the psychoanalytic situation,” in *Methods of Research in Psychotherapy*, eds L. A. Gottschalk and A. H. Auerbach (New York, NY: Appleton-Century-Crofts), 189–208.
- Brumberg, J., and Gumz, A. (2012). Transference interpretations and how they work: a systematic review. *Z. Psychosom. Med. Psychother.* 58, 219–235. doi: 10.13109/zptm.2012.58.3.219
- Brunink, S. A., and Schroeder, H. E. (1979). Verbal therapeutic behavior of expert psychoanalytically oriented, gestalt, and behavior therapists. *J. Consult. Clin. Psychol.* 47, 567–574. doi: 10.1037/0022-006X.47.3.567
- Bucci, W., and Maski, B. (2007). Beneath the surface of the therapeutic interaction: the psychoanalytic method in modern dress. *J. Am. Psychoanal. Assoc.* 55, 1355–1397. doi: 10.1177/000306510705500412
- Butler, S. F., Henry, W. P., and Strupp, H. H. (1995). Measuring adherence in time-limited dynamic psychotherapy. *Psychotherapy* 32, 629–638. doi: 10.1037/0033-3204.32.4.629
- Carkhuff, R. R. (1969). *Helping and Human Relations*. New York, NY: Holt, Rinehart & Winston.
- Carroll, K. M., Nich, C., Sifry, R. L., Nuro, K. F., Frankforter, T. L., Ball, S. A., et al. (2000). A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug Alcohol Depend.* 57, 225–238. doi: 10.1016/S0376-8716(99)00049-6
- Castonguay, L. G. (2011). Psychotherapy, psychopathology, research and practice: pathways of connections and integration. *Psychother. Res.* 21, 125–140. doi: 10.1080/10503307.2011.563250
- Castonguay, L. G. (2013). Psychotherapy outcome: an issue worth Re-Visiting 50 Years Later. *Psychotherapy* 50, 52–67. doi: 10.1037/a0030898
- Castonguay, L. G., and Beutler, L. E. (2006). Principles of therapeutic change: a task force on participants, relationships, and techniques factors. *J. Clin. Psychol.* 62, 631–638. doi: 10.1002/jclp.20256
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 20, 37–46. doi: 10.1177/001316446002000104
- Cohen, J. (1968). Weighted kappa - nominal scale agreement with provision for scaled disagreement or partial credit. *Psychol. Bull.* 70, 213–220. doi: 10.1037/h0026256
- Connolly, M. B., Crits-Christoph, P., Shappell, S., Barber, J., Luborsky, L., and Shaffer, C. (1999). Relation of transference interpretations to outcome in the early sessions of brief supportive-expressive psychotherapy. *Psychother. Res.* 9, 485–495.
- Connolly, M. B., Crits-Christoph, P., Shappell, S., Barber, J. P., and Luborsky, L. (1998). Therapist interventions in early sessions of brief supportive-expressive psychotherapy for depression. *J. Psychother. Pract. Res.* 7, 290–300.
- Connolly Gibbons, M. B., Crits-Christoph, P., Levinson, J., Gladis, M., Siqueland, L., Barber, J. P., et al. (2002). Therapist interventions in the interpersonal and cognitive therapy sessions of the Treatment of Depression Collaborative Research Program. *Am. J. Psychother.* 56, 3–26.
- Connolly Gibbons, M. B., Crits-Christoph, P., Barber, J., Stirman, S. W., Gallop, J., and Goldstein, L. A. (2009). Unique and common mechanisms of change across cognitive and dynamic psychotherapies. *J. Consult. Clin. Psychol.* 77, 801–813. doi: 10.1037/a0016596
- Cooper, M., and Bond, S. (1992). *The Psychodynamic Intervention Rating Scale*. San Francisco, CA: University of California.
- DeRubeis, R. J., Hollon, S. D., Evans, M. D., and Bemis, K. M. (1982). Can psychotherapies for depression be discriminated? A systematic investigation of cognitive therapy and interpersonal therapy. *J. Consult. Clin. Psychol.* 50, 744–756. doi: 10.1037/0022-006X.50.5.744
- Elliott, R. (1979). How clients perceive helper behaviors. *J. Couns. Psychol.* 26, 285–294. doi: 10.1037/0022-0167.26.4.285
- Elliott, R., Barker, C., Caskey, N., and Pistrang, N. (1982). Differential helpfulness of counselor verbal response modes. *J. Consult. Psychol.* 29, 354–361. doi: 10.1037/0022-0167.29.4.354
- Elliott, R., Stiles, W. B., Mahrer, A. R., Hill, C. E., Friedlander, M. L., and Margison, F. R. (1987). Primary therapist response-modes - comparison of 6 rating systems. *J. Consult. Clin. Psychol.* 55, 218–223. doi: 10.1037/0022-006X.55.2.218
- Evans, M. D., Piasecki, J. M., Kriss, M. R., and Hollon, S. D. (1984). *Rater's Manual for the Collaborative Study Psychotherapy Rating Scale - Form 6*. Minneapolis, MN: University of Minnesota and the St. Paul-Ramsey Medical Center.
- Fiedler, F. E. (1950). The concept of the ideal therapeutic relationship. *J. Consult. Psychol.* 14, 239–245. doi: 10.1037/h0058122
- Fiedler, F. E. (1951). Factor analyses of psychoanalytic, nondirective, and Adlerian therapeutic relationships. *J. Consult. Psychol.* 15, 32–38. doi: 10.1037/h0054601
- Finn, R. H. (1970). A note on estimating reliability of categorical data. *Educ. Psychol. Meas.* 30, 71–76. doi: 10.1177/001316447003000106
- Finn, R. H. (1972). Effects of some variations in rating scale characteristics on the means and reliabilities of ratings. *Educ. Psychol. Meas.* 32, 255–265. doi: 10.1177/001316447203200203
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychol. Bull.* 76, 378–382. doi: 10.1037/h0031619
- Fonagy, P. (ed.). (2000). “Section B: reflections on psychoanalytic research problems – an anglo-saxon view,” in *An Open Door Review of Outcome Studies in Psychoanalysis, 2nd Edn* (London: International Psychoanalytic Association), 20–55.
- Forgatch, M. S., and Chamberlain, P. (1982). *The Therapist Behavior Code*. Unpublished instrument and technical report, Eugene, OR: Oregon Social Learning Center.
- Friedlander, M. L. (1982). Counseling discourse as a speech event: revision and extension of the Hill Counselor Verbal Response Category System. *J. Couns. Psychol.* 29, 425–429. doi: 10.1037/0022-0167.29.4.425
- Gaston, L. (1988). *Manual for the Inventory of Therapeutic Strategies - ITS*. Unpublished manuscript. Montreal, QC: McGill University, Department of Psychiatry.
- Gaston, L., and Ring, J. M. (1992). Preliminary results on the inventory of therapeutic strategies. *J. Psychother. Pract. Res.* 1, 135–146.

- Gedo, P. M., and Schaffer, N. D. (1989). An empirical approach to studying psychoanalytic process. *Psychoanal. Psychol.* 6, 277–291. doi: 10.1037/h0085124
- Geißler, P. (2005). *Nonverbale Interaktion in der Psychotherapie Forschung und Praxis im Prozess*. Gießen: Psychosozial-Verlag.
- Gibbons, M. B., Crits-Christoph, P., Barber, J. P., Wiltsey Stirman, S., Gallop, R., Goldstein Elizabeth, A., et al. (2009). Unique and common mechanisms of change across cognitive and dynamic psychotherapies. *J. Consult. Clin. Psychol.* 77, 801–813. doi: 10.1037/a0016596
- Goldfried, M. R., Hayes, A. M., Castonguay, L. G., Drozd, J. E., and Shapiro, D. A. (1997). A comparative analysis of the therapeutic focus in cognitive-behavioral and psychodynamic-interpersonal sessions. *J. Consult. Clin. Psychol.* 65, 740–748. doi: 10.1037/0022-006X.65.5.740
- Goldfried, M. R., Newman, C. F., and Hayes, A. M. (1989). *The Coding System of Therapeutic Focus*. Unpublished manuscript. New York, NY: State University of New York at Stony Brook.
- Goldsamt, L. A., Goldfried, M. R., Hayes, A. M., and Kerr, S. (1992). Beck, Meichenbaum, and Strupp: a comparison of three therapies on the dimension of therapist feedback. *Psychotherapy* 29, 167–176. doi: 10.1037/0033-3204.29.2.167
- Gomes-Schwartz, B., and Schwartz, J. M. (1978). Psychotherapy process variables distinguishing the “inherently helpful” person from the professional psychotherapist. *J. Consult. Clin. Psychol.* 46, 196–197. doi: 10.1037/0022-006X.46.1.196
- Greenhalgh, T., and Peacock, R. (2005). Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. *Br. Med. J.* 331, 1064–1065. doi: 10.1136/bmj.38636.593461.68
- Gumz, A., Horstkotte, J. K., and Kästner, D. (2014). The techniques of the psychodynamic therapist. Types of verbal interventions from a theoretical and practice-oriented perspective. *Z. Psychosom. Med. Psychother.* 60, 219–237. doi: 10.13109/zptm.2014.60.3.219
- Gumz, A. M., Geyer, M., and Brähler, E. (2013). Psychodynamic therapy from the perspective of self-organization. a concept of change and a methodological approach for empirical examination. *Clin. Psychol. Psychother.* 21, 299–310. doi: 10.1002/cpp.1840
- Hahlweg, K., Revenstorf, D., and Schindler, L. (1984). Effects of behavioral marital therapy on couples' communication and problem-solving skills. *J. Consult. Clin. Psychol.* 52, 553–566. doi: 10.1037/0022-006X.52.4.553
- Hannöver, W., Blaut, C., Kniehase, C., Martin, T., and Hannich, H.-J. (2013). Interobserver agreement of the German translation of the Motivational Interviewing Sequential Code for Observing Process Exchanges (MI-SCOPE;D). *Psychol. Add. Behav.* 27, 1196–1200. doi: 10.1037/a0033041
- Heaton, K., Hill, C. E., and Edwards, L. (2010). Comparing molecular and molar methods of judging therapist techniques. *Psychother. Res.* 5, 141–153. doi: 10.1080/10503309512331331266
- Heaton, K. J., Hill, C. E., and Edwards, L. A. (1995). Comparing molecular and molar methods of judging therapist techniques. *Psychother. Res.* 5, 141–153. doi: 10.1080/10503309512331331266
- Hepner, K. A., Azocar, F., Greenwood, G. L., Miranda, J., and Burnam, M. A. (2010). Development of a clinician report measure to assess psychotherapy for depression in usual care settings. *Adm. Policy Ment. Health* 37, 221–229. doi: 10.1007/s10488-009-0249-4
- Hersoug, A. G., Bøgwald, K.-P., and Høglend, P. (2005). Changes of defensive functioning. Does interpretation contribute to change? *Clin. Psychol. Psychother.* 12, 288–296. doi: 10.1002/cpp.444
- Hess, S. A., Knox, S., and Hill, C. E. (2006). Teaching graduate trainees how to manage client anger: a comparison of three types of training. *Psychother. Res.* 16, 282–292. doi: 10.1080/10503300500264838
- Hill, C. E. (1978). Development of a counselor verbal response category system. *J. Consult. Clin. Psychol.* 25, 461–468.
- Hill, C. E., Greenwald, C., Reed, K. G., Charles, D., O'Farrell, M. K., and Carter, J. A. (1981). *Manual for the Counselor and Client Verbal Response Category System*. Columbus, OH: Marathon.
- Hill, C. E., and O'Brien, K. M. (1999). *Helping Skills: Facilitating Exploration, Insight, and Action*. Washington, DC: American Psychological Association.
- Hill, C. E., O'Grady, K. E., and Elkin, I. (1992). Applying the Collaborative study psychotherapy rating scale to rate therapist adherence in cognitive-behavior therapy, interpersonal therapy, and clinical management. *J. Consult. Clin. Psychol.* 60, 73–79. doi: 10.1037/0022-006X.60.1.73
- Hilsenroth, M. J., Bonge, D. R., Blagys, M. D., Ackerman, S. J., and Blais, M. A. (2005). Measuring psychodynamic-interpersonal and cognitive-behavioral techniques: development of the comparative psychotherapy process scale. *Psychother. Theory Res. Pract. Train.* 42, 340–356. doi: 10.1037/0033-3204.42.3.340
- Høglend, P. (2004). Analysis of transference in psychodynamic psychotherapy: a review of empirical research. *Can. J. Psychoanal.* 12, 280–300.
- Hollon, S. D., Waskow, I. E., Evans, M., and Lowery, H. A. (1984). *System for Rating Therapies for Depression*. Los Angeles, CA: Paper presented at the annual meeting of the American Psychiatric Association.
- Holzman, M. S., and Forman, V. P. (1966). A multidimensional content-analysis system applied to the analysis of therapeutic technique in psychotherapy with schizophrenic patients. *Psychol. Bull.* 66, 263–281. doi: 10.1037/h0023825
- Hoyt, M. F., Marmar, C. R., Horowitz, M. J., and Alvarez, W. F. (1981). The therapist action scale and the patient action scale: instruments for the assessment of activities during dynamic psychotherapy. *Psychother. Theory Res. Pract.* 18, 109–116. doi: 10.1037/h0085950
- Jones, E. E. (1985). *Manual for the Psychotherapy Process Q-set*. Berkeley, CA: University of California.
- Jones, E. E., Krupnick, J. L., and Kerig, P. K. (1987). Some gender effects in a brief psychotherapy. *Psychotherapy* 24, 336–352. doi: 10.1037/h0085724
- Jones, E. E., and Windholz, M. (1990). The psychoanalytic case study: toward a method for systematic inquiry. *J. Am. Psychoanal. Assoc.* 38, 985–1015. doi: 10.1177/000306519003800405
- Karl, N. J., and Abeles, N. (1969). Psychotherapy process as a function of the time segment sampled. *J. Consult. Clin. Psychol.* 33, 207–212. doi: 10.1037/h0027173
- Kerr, S., Goldfried, M. R., Hayes, A. M., Castonguay, L. G., and Goldsamt, L. A. (1992). Interpersonal and intrapersonal focus in cognitive-behavioral and psychodynamic-interpersonal therapies: a preliminary analysis of the Sheffield Project. *Psychother. Res.* 2, 266–276. doi: 10.1080/1050330921233133024
- Koenigsberg, H. W., Kernberg, O. F., Haas, G., Lotterman, A., Rockland, L., and Selzer, M. (1985). Development of a scale for measuring techniques in the psychotherapy of borderline patients. *J. Nerv. Ment. Dis.* 173, 424–431. doi: 10.1097/00005053-198507000-00006
- Koenigsberg, H. W., Kernberg, O. F., Rockland, L., Appelbaum, A., Carr, A., and Kernberg, P. (1988). “Developing an instrument for characterizing psychotherapy techniques in studies of the psychotherapy of borderline patients,” in *Psychoanalytic Process Research Strategies*, eds H. Dahl, H. Kächele and H. Thomä (New York: Springer-Verlag), 147–62.
- Lambert, M. J. (2013). “The efficacy and effectiveness of psychotherapy,” in *Bergin and Garfield's Handbook of Psychotherapy and Behavior Change*, Vol. 6, ed M. J. Lambert (New Jersey: NJ; Wiley), 169–218.
- Liese, B. S., Shepherd, D. D., Cameron, C. L., and Ojeleye, A. E. (1995). Teaching psychological knowledge and skills to family physicians. *J. Clin. Psychol. Med. Settings* 2, 21–38. doi: 10.1007/BF01988625
- Light, R. J. (1971). Measures of response agreement for qualitative data: some generalizations and alternatives. *Psychol. Bull.* 76, 365–377. doi: 10.1037/h0031643
- Luborsky, L. (1984). *Principles of Psychoanalytic Psychotherapy: A Manual for Supportive-Expressive (SE) Treatment*. New York, NY: Basic Books.
- Luborsky, L., Bachrach, H., Graff, H., Pulver, S., and Christoph, P. (1979). Preconditions and consequences of transference interpretations. *J. Nerv. Ment. Dis.* 167, 391–401. doi: 10.1097/00005053-197907000-00001
- Madonik, B. G. (2001). *I Hear What You Say, but What Are You Telling Me? The Strategic Use of Nonverbal Communication in Mediation*. San Francisco, CA: Jossey-Bass.
- Malan, D. H. (1963). *A Study of Brief Psychotherapy*. London: Tavistock.
- Malan, D. H. (1976). *Toward the validation of dynamic psychotherapy*. New York, NY: Plenum Press.
- Margison, F., Barkham, M., Evans, C., McGrath, G., Mellor-Clark, J., Audin, K., et al. (2000). Measurement and psychotherapy: evidence based

- practice and practice based evidence. *Br. J. Psychiatry* 177, 123–130. doi: 10.1192/bjp.177.2.123
- Mark, D., and Faude, J. (1995). “Supportive-expressive therapy for cocaine abuse,” in *Dynamic Therapies for Psychiatric Disorders (Axis I)*, eds J. P. Barber and P. Crits-Christoph (New York, NY: Basic Books), 294–331.
- Marziali, E. A., and Sullivan, J. M. (1980). Methodological issues in the content analysis of brief psychotherapy. *Br. J. Med. Psychol.* 53, 19–27. doi: 10.1111/j.2044-8341.1980.tb02864.x
- McCarthy, K. S., and Barber, J. P. (2009). The Multitheoretical List of Therapeutic Interventions (MULTI): initial report. *Psychother. Res.* 19, 96–113. doi: 10.1080/10503300802524343
- McCullough, L. (1988). Psychotherapy interaction coding system manual: the PIC system. *Soc. Behav. Sci. Doc.* 18.
- Mercer, D., and Woddy, G. (1992). *Addiction Counseling*. University of Pennsylvania: VAMC Center for Studies of Addiction.
- Mergenthaler, E. (2015). “Resonating minds: a school-independent theoretical conception and its empirical application to psychotherapeutic processes, in *Visions in Psychotherapy Research and Practice: Reflections from Presidents of the Society for Psychotherapy Research*, eds B. Strauss, J. Barber and L. Castonguay (New York, NY: Taylor & Francis), 292–314.
- Milbrath, C., Bond, M., Cooper, S., Znoj, H. J., Horowitz, M. J., and Perry, J. C. (1999). Sequential consequences of therapists’ interventions. *J. Psychother. Pract. Res.* 8, 40–54.
- Miller, W. R., Moyers, T. B., Ernst, D., and Amrhein, P. (2008). *Manual for the Motivational Interviewing Skill Code (MISC)*. The University of New Mexico; Center on Alcoholism, Substance Abuse, and Addictions.
- Moyers, T. B., and Martin, T. (2006). Therapist influence on client language during motivational interviewing sessions. *J. Subst. Abuse Treat.* 30, 245–251. doi: 10.1016/j.jsat.2005.12.003
- Moyers, T., Martin, T., Catley, D., Harris, K. J., and Ahluwalia, J. S. (2003). Assessing the integrity of motivational interviewing interventions: reliability of the Motivational Interviewing Skills Code. *Behav. Cogn. Psychother.* 31, 177–184. doi: 10.1017/S1352465803002054
- Norcross, J. C. (2011). (Ed.). *Psychotherapy Relationships that Work: Evidence-based Responsiveness*, 2nd Edn. New York, NY: Oxford University Press.
- Ogrodniczuk, J. S., and Piper, W. E. (1999). Measuring therapist technique in psychodynamic psychotherapies development and use of a new scale. *J. Psychother. Pract. Res.* 8, 142–154.
- O’Malley, S. S., Suh, C. S., and Strupp, H. H. (1983). The Vanderbilt psychotherapy process scale: a report on the scale development and a process-outcome study. *J. Consult. Clin. Psychol.* 51, 581–586. doi: 10.1037/0022-006X.51.4.581
- Orlinsky, D. E., and Howard, K. I. (1967). The good therapy hour: experiential correlates of patients’ and therapists’ evaluations of therapy sessions. *Arch. Gen. Psychiatry* 16, 621–632. doi: 10.1001/archpsyc.1967.01730230105013
- Orlinsky, D. E., and Howard, K. I. (1975). *Varieties of Psychotherapeutic Experience: Multivariate Analyses of Patients’ and Therapists’ Reports*. New York, NY: Teachers College Press.
- Orlinsky, D. E., and Howard, K. I. (1987). A generic model of psychotherapy. *J. Integr. Eclectic Psychother.* 6, 6–27.
- Paniagua, C. (2003). Problems with the concept ‘interpretation’. *Int. J. Psychoanal.* 84, 1105–1123. doi: 10.1516/17U4-KBU0-NFNB-2E8K
- Pawelczyk, J. (2011). *Talk as Therapy: Psychotherapy in a Linguistic Perspective*. Berlin; Boston: Mouton de Gruyter.
- Pfammatter, M., and Tschacher, W. (2012). Wirkfaktoren der Psychotherapie – eine Übersicht und Standortbestimmung. *Z. Psychiatr. Psychol. Psychother.* 60, 67–76. doi: 10.1024/1661-4747/a000099
- Piper, W. E., Debbane, E. G., de Carufel, F. L., and Bienvenu, J. P. (1987). A system for differentiating therapist interpretations from other interventions. *Bull. Menninger Clin.* 51, 532–550.
- Piper, W. E., Joyce, A. S., McCallum, M., and Azim, H. F. (1998). Interpretive and supportive forms of psychotherapy and patient personality variables. *J. Consult. Clin. Psychol.* 66, 558–567. doi: 10.1037/0022-006X.66.3.558
- Samoilov, A., Goldfried, M. R., and Shapiro, D. (2000). Coding system of therapeutic focus on action and insight. *J. Consult. Clin. Psychol.* 68, 513–514. doi: 10.1037/0022-006X.68.3.513
- Schindler, L., Hohenberger-Sieber, E., and Hahlweg, K. (1989). Observing client-therapist interaction in behaviour therapy: development and first application of an observational system. *Br. J. Clin. Psychol.* 28, 213–226.
- Schreier, M. (2012). *Qualitative Content Analysis in Practice*. Los Angeles, CA; London: Sage.
- Shapiro, D. A., Barkham, D. A., and Irving, D. L. (1980). *A Modified Helper Behavior Rating Manual (Social and Applied Psychology Unit Memo No. 415)*. Sheffield: Department of Psychology; University of Sheffield.
- Shapiro, D. A., Barkham, M., and Irving, D. L. (1984). The reliability of a modified Helper Behavior Rating System. *Br. J. Med. Psychol.* 57, 45–48. doi: 10.1111/j.2044-8341.1984.tb01579.x
- Shapiro, D. A., and Startup, M. J. (1990). *Raters’ Manual for the Sheffield Psychotherapy Rating Scale (Memo No. 1154)*. Sheffield: University of Sheffield; MRC/ESRC Social and Applied Psychology Unit.
- Shrout, P., and Fleiss, J. L. (1979). Intraclass correlation: uses in assessing rater reliability. *Psychol. Bull.* 86, 420–428. doi: 10.1037/0033-2909.86.2.420
- Silberschatz, G., Fretter, P. B., and Curtis, J. T. (1986). How do interpretations influence the process of psychotherapy? *J. Consult. Clin. Psychol.* 54, 646–652. doi: 10.1037/0022-006X.54.5.646
- Snyder, W. U. (1945). An investigation of the nature of non-directive psychotherapy. *J. Gen. Psychol.* 33, 193–223. doi: 10.1080/00221309.1945.10544506
- Startup, M., and Shapiro, D. A. (1993). Dimensions of cognitive therapy for depression: a confirmatory analysis of session ratings. *Cognit. Ther. Res.* 17, 139–151. doi: 10.1007/BF01172962
- Stiles, W. B. (1978). *Manual for a Taxonomy of Verbal Response Modes*. Chapel Hill: University of North Carolina at Chapel Hill; Institute for Research in Social Science.
- Stiles, W. B. (1992). *Describing Talk: A Taxonomy of Verbal Response Modes*. Newbury Park: Sage.
- Stiles, W. B., Hardy, G. E., Barkham, M., Rees, A., Shapiro, D. A., Startup, M., et al. (1996). Therapist session intentions in cognitive-behavioral and psychodynamic-interpersonal psychotherapy. *J. Couns. Psychol.* 43, 402–414. doi: 10.1037/0022-0167.43.4.402
- Stiles, W. B., Shapiro, D. A., and Firthcozens, J. A. (1989). Therapist differences in the use of verbal response mode forms and intents. *Psychotherapy* 26, 314–322. doi: 10.1037/h0085441
- Strauss, B., Strupp, H. H., Burgmeier-Lohse, M., Wille, H., and Storm, S. (1992). Deutschsprachige Version der Vanderbilt-Psychotherapie-Skalen: beschreibung und Anwendung in zwei Kurztherapien. *Z. Klin. Psychol. Psychopathol. Psychother.* 40, 411–430.
- Strupp, H. H. (1957). A multidimensional system for analyzing psychotherapeutic techniques. *Psychiatry* 20, 293–312.
- Strupp, H. H. (1966). *A Multidimensional System für Analyzing Psychotherapeutic Communications: Manual, 2nd Edn*. Chapel Hill: University of North Carolina.
- Strupp, H. H., Horowitz, L. M., and Lambert, M. J. (1997). *Measuring Patient Changes in Mood, Anxiety, and Personality Disorders. Toward a Core Battery*. Washington, DC: American Psychological Association.
- Tinsley, H. E. A., and Weiss, D. J. (1975). Interrater reliability and agreement of subjective judgments. *Res. Methodol.* 22, 358–376. doi: 10.1037/h0076640
- Town, J. M., Hardy, G. E., McCullough, L., and Stride, C. (2012a). Patient affect experiencing following therapist interventions in short-term dynamic psychotherapy. *Psychother. Res.* 22, 208–219. doi: 10.1080/10503307.2011.637243
- Town, J. M., McCullough, L., and Hardy, G. E. (2012b). Therapist interventions in short-term dynamic psychotherapy: six expert treatments. *Br. J. Guid. Couns.* 40, 31–42. doi: 10.1080/03069885.2011.621518
- Tschacher, W., Junghan, U. M., and Pfammatter, M. (2014). Towards a taxonomy of common factors in psychotherapy - results of an expert survey. *Clin. Psychol. Psychother.* 21, 82–96. doi: 10.1002/cpp.1822
- Tversky, A., and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science* 185, 1124–1131. doi: 10.1126/science.185.4157.1124
- Vaughan, S. C., Spitzer, R., Davies, M., and Roose, S. (1997). The definition and assessment of analytic process: can analysts agree? *Int. J. Psychoanal.* 78, 959–973.
- Waldron, S., Scharf, R., Crouse, J., Firestein, S., Burton, A., and Hurst, D. (2004a). Saying the right thing at the right time: a view through the lens

- of the Analytic Process Scales (APS). *Psychoanal. Q.* 73, 1079–1125. doi: 10.1002/j.2167-4086.2004.tb00193.x
- Waldron, S., Scharf, R. D., Hurst, D., Firestein, S. K., and Burton, A. (2004b). What happens in a psychoanalysis? A view through the lens of the analytic process scale (APS). *Int. J. Psychoanal.* 85, 443–466. doi: 10.1516/5PPV-Q9WL-JKA9-DRCK
- Wampold, B. E. (2001). *The Great Psychotherapy Debate: Models, Methods, and Findings*. Hillsdale, NJ: Erlbaum.
- Wampold, B. E., and Imel, Z. E. (2015). *The Great Psychotherapy Debate: The Evidence for What Makes Psychotherapy Work*. New York, NY: Routledge.
- Winston, A., Pollack, J., McCullough, L., Flegenheimer, W., Kestenbaum, R., and Trujillo, M. (1991). Brief psychotherapy of personality disorders. *J. Nerv. Ment. Dis.* 179, 188–193. doi: 10.1097/00005053-199104000-00002
- Young, J. E., and Beck, A. T. (1980). *Cognitive Therapy Scale: Rating Manual*. Philadelphia, PA: University of Pennsylvania.
- Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2015 Gumz, Treese, Marx, Strauss and Wendt. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.