Psychotherapy Research

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tpsr20

The therapeutic effects of the therapists' ability to pass their patients' tests in psychotherapy

Ramona Fimiani, Francesco Gazzillo, Bernard Gorman, Jessica Leonardi, Giuseppe Stefano Biuso, Martina Rodomonti, Camilla Mannocchi & Federica Genova

To cite this article: Ramona Fimiani, Francesco Gazzillo, Bernard Gorman, Jessica Leonardi, Giuseppe Stefano Biuso, Martina Rodomonti, Camilla Mannocchi & Federica Genova (2022): The therapeutic effects of the therapists' ability to pass their patients' tests in psychotherapy, Psychotherapy Research, DOI: 10.1080/10503307.2022.2157227

To link to this article: https://doi.org/10.1080/10503307.2022.2157227

Published online: 27 Dec 2022.

Psychotherapy Research, 2022
https://doi.org/10.1080/10503307.2022.2157227

EMPIRICAL PAPER

The therapeutic effects of the therapists' ability to pass their patients' tests in psychotherapy

RAMONA FIMIANI[1], FRANCESCO GAZZILLO [1], BERNARD GORMAN[2], JESSICA LEONARDI [1], GIUSEPPE STEFANO BIUSO[3], MARTINA RODOMONTI[3], CAMILLA MANNOCCHI[3], & FEDERICA GENOVA[4]

[1]*Department of Dynamic and Clinical Psychology, and Health Studies, "Sapienza" University of Rome, Rome, Italy;* [2]*Derner Institute of Advanced Psychology Studies, Adelphi University, Long Island, NY, USA;* [3]*Control Mastery Theory Italian Group, Rome, Italy &* [4]*Alma Mater Studiorum University of Bologna, Bologna, Italy*

(Received 17 October 2022; revised 6 December 2022; accepted 7 December 2022)

Abstract
Introduction According to Control-Mastery Theory (CMT)—a cognitive-dynamic relational theory of mental functioning, psychopathology, and psychotherapy—patients come to therapy with an unconscious plan to disprove their pathogenic beliefs and achieve adaptive goals. One of the primary ways patients work to disconfirm their pathogenic beliefs is by testing them within the therapeutic relationship. **Objectives:** The present study aimed to replicate and expand the results of previous studies suggesting that therapists' responses that disconfirmed patient's pathogenic beliefs were predictive of patients' within-session progress. Moreover, we wanted to investigate whether these interventions correlated with the therapeutic alliance. **Methods:** Transcriptions of 81 sessions from five brief psychodynamic psychotherapies were assessed by 11 independent raters. For each case, the patient's plan was formulated and tests identified, the accuracy of the therapist's responses to these tests was rated, and the impact of the therapist's interventions on the patient's subsequent communications and their relationship with the therapeutic alliance was measured. **Results:** The results supported the central hypothesis of the CMT that when the therapist's interventions passed the patient's tests, the patient showed signs of improvement. Moreover, the ability of the therapist to pass the patient's tests correlated with the therapeutic alliance. **Conclusions**: The clinical implications and the limitations of these findings are discussed, together with the relevance of a good case formulation for clinicians' optimal responsiveness.

Keywords: Control Mastery Theory; process research; patient's plan; testing activity; therapeutic alliance

Clinical or methodological significance of this article: This study shows the relevance of a good case formulation for clinicians' responsiveness, and the importance of clinicians' ability to disprove their patients' pathogenic beliefs for the therapeutic progress. When therapists are able to disprove their patients' pathogenic beliefs by passing their tests, patients become more able to understand themselves and address their problems, and become more involved in their therapies.

Research in psychotherapy has clearly shown that psychotherapy is effective (APA, 2012; Barkham & Lambert, 2021; Cuijpers et al., 2021; Dragioti et al., 2017) and that, if certain conditions are respected – the therapy is a bona fide treatment guided by a manual and based on a psychological theory of psychopathology and therapeutic process; therapists believe that the treatments they are delivering are effective; therapists are appropriately trained and so on (Laska et al., 2014; Frank & Frank, 1991) - all the therapies aimed at treating a specific disorder are equally effective, independently from the theory they are based on (Steinert et al., 2017; Wampold & Imel, 2015; Wampold et al., 1997).

Correspondence concerning this article should be addressed to Francesco Gazzillo, Department of Dynamic and Clinical Psychology, and Health Studies, "Sapienza" University of Rome, Rome, Italy. Email: francesco.gazzillo@uniroma1.it

This last finding, together with the data that shows how only a small amount of outcome variance is explained by the use of specific therapeutic techniques (Ahn & Wampold, 2001; Lambert & Barley, 2001), has pointed out the importance of non-specific and relational factors in understanding the efficacy of psychotherapy (Norcross & Lambert, 2019; Castonguay et al., 2019; Wampold, 2015). The available research clearly demonstrates the healing capacity of the therapy relationship and the beneficial value of adapting that relationship to individual patients beyond their diagnosis. Among the factors whose relevance has been more convincingly shown we find: therapeutic alliance; the perceived empathy of the therapist; positive regard from the therapist; therapist's and patient's congruence or genuineness; positive expectations in the patient; the ability of therapists to adjust their approach to specific features of their patients, such as their level of reactance or their spiritual values, etc. (for an overview of these factors, see Norcross & Lambert, 2019; Norcross & Wampold, 2019). Moreover, many authors have recently become dissatisfied with the effectiveness of empirically supported treatments for specific disorders (for some of their limitations, see Westen et al., 2004; Shedler, 2018), and they have suggested how it would be probably more fruitful to pass from disorder-specific treatments to *case-specific treatments* (Critchfield et al., 2022; Kramer, 2019, 2020), which implies giving more attention to the elaboration of reliable case formulations that give therapeutic indications useful for optimizing the effectiveness of a treatment (see, for example, Barber & Solomonov, 2019; Zilcha-Mano, 2021).

Finally, one of the constructs that are receiving more attention in contemporary literature on psychotherapy research is the construct of *optimal responsiveness* (Kramer, 2021; Stiles, 2021; Silberschatz, 2021; Wu & Levitt, 2020; Kramer & Stiles, 2015; Bacal, 1998; Stiles et al., 1998), which stress how therapists should not only be able to provide case-specific treatments, but also to "adapt" their approaches to the moment to moment shifts in patients' needs. In fact, an optimal responsiveness of clinicians appears to be a *sine qua non* for a psychotherapy to be a source of corrective emotional experiences (Alexander & French, 1946; Castonguay & Hill, 2012) for patients, i.e., experiences that contradict their negative expectations connected to their traumas and adverse experiences.

In our opinion, one of the theoretical models which is better suited both to help therapists develop case formulations that are reliable and relevant for therapy outcome, and to be optimally attuned with their patients' moment-to-moment needs, is the Control-Mastery Theory (CMT; Gazzillo, 2021; Silberschatz, 2005; Weiss, 1993; Weiss et al., 1986).

The Basic Concepts of Control-Mastery Theory

CMT is an integrative, cognitive-dynamic relational theory of mental functioning, psychopathology, and psychotherapeutic processes. It is based on the following assumptions: that the overarching aim of mental functioning is to adapt to reality; that patients are unconsciously motivated to solve their problems and master their traumas; that human beings can unconsciously perform many of the complex mental activities normally connected to consciousness (such as developing and testing beliefs, goals and plans and assessing reality); and that the main regulating principle of mental functioning is one of safety. Most of these assumptions have been confirmed by recent developments in cognitive psychology, developmental psychology, moral psychology, and evolutionary psychology (see Leonardi et al., 2022; Gazzillo et al., 2020).

According to CMT, functional psychopathology derives from pathogenic beliefs generally developed during childhood and adolescence to adapt to traumas (Aafjes-van Doorn et al., 2020; Fimiani et al., 2020; Silberschatz & Aafjes-van Doorn, 2017). Pathogenic beliefs associate the pursuit of healthy and adaptive goals with dangers to the self, important others, and important relationships (Gazzillo, 2022; Weiss, 1997). Given that these beliefs are grim and constricting, patients are highly motivated to become conscious of them, disprove them, and master the traumas that lie beneath them. They can then pursue their adaptive goals and overcome their problems (Weiss, 1990).

One of the main ways patients try to disprove their pathogenic beliefs is by testing them in their relationships with the therapist. Tests include communications, attitudes, and behaviors consciously or unconsciously aimed at disproving pathogenic beliefs (Gazzillo et al., 2019). To pass a patient's tests, the therapist has to respond to them in a way that is felt by the patient to be a disconfirmation of the pathogenic belief being tested (Novak et al., 2022). The therapist may hypothesize that a patient is testing them when the patient expresses emotions that are stronger than usual, when the therapist feels pressure to intervene in some way, when the patient makes an implicit or explicit request, or when the patient behaves more absurdly, illogically, or provocatively than usual (Weiss, 1993).

According to CMT, when the therapist passes a patient's tests and thus contributes to the disconfirmation of the patient's pathogenic beliefs, the patient is likely to become less anxious and depressed and more involved in the therapeutic work and in the therapeutic relationship. Moreover, the patient may gain new insights, become more elaborative, become bolder in dealing with their problems, and show other signs of improvement. This is because disconfirming a pathogenic belief means reducing patients' negative expectations and the negative emotions of fear, anxiety and depression connected to them. Moreover, it means supporting patients' sense of safety and their hope that, even with the help of the therapist, they can pursue their goals.

Let us consider a simple example of a patient who, because he was severely neglected by his parents during his childhood, developed the pathogenic belief that he deserves rejection and consequently is afraid of being rejected in virtually all the situations that are emotionally relevant for him. If this patient unexpectedly threatens to quit the treatment, it is legitimate to consider that he is (unconsciously) testing to see if his therapist wants to "get rid" of him thinking that he deserves rejection (in the hope that the therapist will not make him quit). If the therapist persists and encourages him not to quit, the test is "passed", and the patient will feel safer and become more engaged. By contrast, if the therapist does or says something that suggests the opposite (e.g., going along with the patient's desire to quit prematurely), the therapist will have failed the test because the patient may easily take the therapist' behavior as evidence that the therapist is rejecting him. If the therapist fails such a test, the patient is likely to remain anxious or depressed. He may react less enthusiastically and productively to the therapist's interventions, or not react at all or change the topic. The patient may also try to "*coach*" the therapist about what he needs (Bugas et al., 2021; Bugas & Silberschatz, 2000; Kealy et al., 2022); for example, the patient may react to the fact that therapist accepted his idea of terminating the treatment by talking about his regretting the fact of having interrupted his relationship with a former girlfriend.

Several studies have confirmed these hypotheses about testing. Horowitz et al.'s (1975), 100-hour study of a patient in psychoanalysis showed that when the patient was being tested he became more anxious, but when the therapist passed his tests, he became more relaxed and could more easily recover previously repressed memories. Silberschatz (1986) and Silberschatz et al.'s (1986) studies offered further confirmation, demonstrating that when the therapist passed the patient's tests, the patient became more positive towards other people and more able to work through her feelings. A fourth and final study (Silberschatz & Curtis, 1993) conducted on two time-limited therapies found the same.

According to CMT, the patient comes to therapy with an unconscious plan to get better (Gazzillo et al., 2021, 2022; Weiss, 1994). This plan may be thought of as comprising: (a) the healthy and adaptive goals that the patient wants to pursue; (b) the pathogenic beliefs that the patient wants to disprove that are hindering the pursuit of these goals; (c) the traumas that lead to the development of the patient's pathogenic beliefs and that the patient wants to master; (d) tests the patient may initiate and how they may want the therapist to respond to their test; and (e) information or insights that may be helpful to the patient to understand and overcome pathogenic beliefs. The Plan Formulation Method (PFM; Curtis & Silberschatz, 2022) is an empirically validated procedure developed for a reliable formulation of the plan of patients on the basis of the first two to ten sessions of their treatments. Several studies have confirmed the CMT hypothesis that interventions supporting the patient's plan lead to immediate and long-term improvements (Silberschatz, 2017; Silberschatz & Curtis, 1986).

A good formulation of a patient's plan for psychotherapy can guide the clinician in understanding which goals that patient want to pursue; which pathogenic beliefs are obstructing them and need to be disproved; which traumas they need to master better; how that patient may test the clinician during the therapy and how the clinician can pass their test; what they patient may want to understand about themselves and their problems. For these reasons, a good enough formulation of the plan of a patient can improve the therapist responsiveness.

Aims and Hypotheses

The present study aimed to replicate and expand upon the findings of previous empirical studies on testing that have shown that the degree to which a therapist passes or fails a patient's tests correlates with the patient's immediate therapeutic progress. Moreover, we wanted to assess the correlation between the therapist's ability to pass patient's tests and the therapeutic alliance within the same session, given that the therapeutic alliance is one of the best predictors of the outcome of psychotherapy (see, for example, Flückiger et al., 2018; Zilcha-Mano, 2017; Xu & Tracey, 2015).

Based on the above, the following hypotheses were proposed:

H1: The therapist's ability to pass the patent's test significantly predicts an improvement in patient's in-session ease and therapeutic productivity as showed by: an increase in the patient's state of freedom, relaxation, and comfort; an improvement in the depth and breadth of patient's emotional understanding of his/her issues and in his/her involvement and collaboration with the therapist; an increase in the patient's ability to flexibly shift to and from experiencing and reflecting; an improvement in the patient's ability to work through his/her problematic patterns of relating and feeling; and an improvement in patient's ability to address his/her central issues and concerns with boldness, ease, and interest.

H2: The average ability of the therapist to pass the patient's tests in a session predicts the patient's average ease and therapeutic productivity following the therapist's interventions within the same session, as an effect of the therapist's repeated disconfirmations of the patient's pathogenic beliefs (this is following Silberschatz & Curtis, 1993).

H3: The degree to which therapist is able to pass the patient's tests correlates with the therapeutic alliance during the same session.

Method

Sample

The opportunity for this study derived from the availability of 39 fully recorded and transcribed brief psychodynamic therapies (16 sessions), all in the collection of the San Francisco Psychotherapy Research Group.[1] This material was collected as part of the Mount Zion Brief Therapy Research Project, a study investigating the process and outcome of brief psychodynamic psychotherapy (Silberschatz et al., 1991).

All patients were self-referred and screened to ensure their suitability for brief treatment. They were required to meet the following minimum acceptance criteria: (a) no evidence of psychosis; (b) no evidence of severe substance abuse; (c) no evidence of organic brain impairment; and (d) no evidence of suicide risk. The predominant conditions were anxiety and depressive disorders. Patients who met the criteria were asked to sign an informed consent form explaining that all interviews and therapy sessions would be audio-recorded, transcribed, and used for research purposes. Strict anonymity and confidentiality were guaranteed by a systematic coding process that deleted all names and other potentially identifying data.

The therapists were experienced psychiatrists and psychologists with a psychodynamic orientation. All had a minimum of 3 years' post-graduate experience and had received training in brief psychodynamic therapy. Therapists received no information about the patients before treatment began other than that they had been screened and accepted for brief therapy.

The data sources for the present study were verbatim transcriptions of five cases (81 sessions). In addition to the 16 therapy sessions, each treatment included an intake interview that was conducted by an independent evaluator different from the treating therapist, which was also audio-recorded and transcribed. The five patients included in this study (two men and three women aged between 20–61 years) were randomly selected from an overall sample of 39 patients in the Mount Zion Brief Therapy Research Project, after excluding two cases used in a previous study (Silberschatz & Curtis, 1993) that investigated the relationship between the therapist's ability to pass the patient's tests and the patient's improvement in the therapeutic process.

Millie, a 54-year-old married woman, sought therapy because she felt depressed. She had lost her income due to her husband's illness; she needed to find employment but felt incapable of holding down a job. In addition, she could not decide on what kind of work to pursue. She also described feeling lonely and alienated from her peers.

Hugo, a 50-year-old man, was a commercial freelance graphic artist living with his girlfriend. He entered therapy because he anticipated that the relationship would be ending and that he would need some support to handle the subsequent loneliness and the difficulties he had feeling close to people.

Myra, a 31-year-old self-employed graphic artist, sought therapy because she was having difficulty breaking off an unsatisfactory relationship with her partner.

Vicky, a 20-year-old student, entered therapy because of sexual problems and her difficulties in becoming emotionally involved with men. The same therapist treated her and Myra.

Finally, Leon, a 61-year-old man, sought therapy because he felt unhappy in his administrative role as a director of a social service agency. Moreover, he complained of difficulties in making decisions and a tendency to procrastinate.

Hugo completed all 16 therapy sessions, Millie and Myra completed 15 out of 16, Vicky completed 18 out of 20, and Leon completed 12 out of 16. The cases reported here were studied after the therapy sessions had come to an end.

Measures

To formulate the patients' plans and assess the impact of the therapists' responses to the patients'

tests on the latter's subsequent communications and the therapeutic alliance, the following measures were used:

Relevance scale. The Relevance Scale (Curtis et al., 1994) is a 5-point Likert scale ranging from 0 (*not relevant*) to 4 (*very highly relevant*). It assesses the relevance to the specific case of each different items of the patient's plan formulation. The inter-rater reliability between two independent judges for each patient's plan formulation was excellent. Across the five cases assessed on the basis of the first two sessions (the intake interview and first session), the overall average ICC was .90, ranging from a minimum of .84 for Leon's plan formulation to a maximum of .97 for Hugo's. Excellent levels of reliability among the independent judges were also obtained for each patient's plan component, ranging from a minimum of .84 (the patient's goals and tests) to a maximum of .93 (the patient's traumas). The data was obtained by averaging the ICC values of each patient's plan component for the five patients' plan formulations.

Patient scale of testing. The Patient Scale of Testing (a revised version of the Patient Scale of Key Test; see Silberschatz et al., 1986) is a 5-point Likert scale ranging from 0 (*the segment is not an example of the patient testing the therapist*) to 4 (*the segment is an excellent, clear-cut example of the patient testing the therapist*). It assesses the degree to which each of the previously selected segments represents a clear example of a test, that is, an instance of the patient testing a pathogenic belief in the therapeutic relationship. The pooled intraclass correlation coefficient (ICC) of the Patient Scale of Testing assessed by three independent judges on 707 segments of patients' communications was .76.

Therapist scale of passing versus failing tests. The Therapist Scale of Passing Versus Failing Tests (Silberschatz, 1986) is a 7-point Likert scale ranging from 0 (*the therapist's response is an explicit, clear-cut example of failing the patient's test*) to 6 (*the therapist's response is an excellent, clear-cut example of passing the patient's test*). It assesses the degree to which the therapist passes or fails the patient's tests, that is, the extent to which the therapist's interventions disconfirm or confirm the patient's pathogenic beliefs. The pooled intraclass correlation coefficient (ICC) of the Therapist Scale of Passing versus Failing Tests assessed by three independent judges on 707 segments of patients' communications was .72.

Analytic process scale. The Analytic Process Scales (APS; Waldron et al., 2004) contains 29 Likert format items (rated 0–4), with each aimed at assessing a dimension of the therapeutic process. The patient's contribution to the therapy is assessed with 12 items, the therapist's contribution with 16 items, and their interaction with one item. The definitions of each of these items and clinical examples at the level of 0 (*absent*), 2 (*present and explicit*), and 4 (*present, explicit, complex, and articulate*) have been assembled into an 81-page coding manual (Scharf et al., 2010). The *overall productivity of the patient's communications* scale used in the present study assesses the degree to which there is a sense of forward movement during the segment in the depth or breadth of the patient's or the rater's emotional understanding, in the intensity of the patient's involvement and collaboration with the therapist, or in the quality of other emotional expressions. The intraclass correlation coefficient (ICC) of the overall productivity of the patient's communications rated by two independent judges on 678 segments of patients' communications was .60.

Dynamic interaction scales. The Dynamic Interaction Scales (DIS; Waldron et al., 2013) consist of 12 Likert-format items (rated on a 0–4 scale similar to the one used for the APS assessment). They are aimed specifically at the assessment of the global, relational, and interactional aspects of the therapeutic process. Five DIS items are used to assess the therapist's contribution, three DIS items are used to assess the patient's contribution, and four DIS items are used to assess their interaction. Definitions of each level of the scales have been assembled into a 7-page coding manual (Waldron, 2011). Two DIS patient items were used in the present study. One item (DIS_7) assesses *the degree to which the patient flexibly shifts to and from experiencing and reflecting;* that is to say, how well the patient seems to move effortlessly between being immersed in associations and/or the current experience in the room with the therapist and reflecting on it. The other item (DIS_9) assesses *how well the patient is working with his/her problematic patterns of relating and feelings*, that is, the degree to which the patient addresses central issues and concerns in a compelling, productive manner. The intraclass correlation coefficient (ICC) of the DIS items assessed by two independent judges on 678 segments of patients' communications was .68 for DIS_7 and .63 for DIS_9.

Boldness scale. The Boldness Scale (Caston et al., 1986) is a 9-point Likert scale ranging from 1

(*the patient manifests clear-cut retreat or inhibition*) to 9 (*the patient manifests a bold and interested tackling of issues and plunges ahead*). It assesses the degree to which the patient pursues a deepening exploration of emotionally relevant personal issues with boldness, ease, and interest. The judge combines two factors in making this rating: (a) if the patient focuses on important rather than trivial issues; and (b) if the patient deal with these issues with ease. The instructions for raters are contained in the appendix of the volume "The psychoanalytic process: Theory, clinical observation, and empirical research" (Weiss et al., 1986). The intraclass correlation coefficient (ICC) of the Boldness Scale assessed by two independent judges on 678 segments of patients' communications was .73.

Relaxation scale. The Relaxation Scale (Curtis et al., 1986) is a 5-point Likert scale ranging from 0 (*the patient seems uncomfortable, beleaguerment, driven, defensive, constricted, tense, tight*) to 4 (*the patient seems relaxed, free, unconstricted, spontaneous relatively more so than in 3*). It measures the patient's psychic state of freedom, relaxation, and comfort versus that of anxiety, drivenness, and beleaguerment. Low scores suggest that patient is tense, stressed, rigid and anxious, or driven. High scores are given when the patient is relaxed, playful, able to experience a wide range of feelings, and able to associate freely. The instructions for raters are contained in the appendix of the volume "The psychoanalytic process: Theory, clinical observation, and empirical research" (Weiss et al., 1986). The intraclass correlation coefficient (ICC) of the Relaxation Scale assessed by two independent judges on 678 segments of patients' communication was .59.

Working alliance inventory – observer form. The Working Alliance Inventory – Observer Form (WAI-O; Darchuk et al., 2000) is an observer measure designed to capture Bordin's pantheoretical perspective of the working alliance. The 36 Likert-format items, ranging from 1 (*very strong evidence against*) to 7 (*very strong evidence*), consist of three subscales: the Goal subscale addresses the extent to which therapy goals are important, mutual, and capable of being accomplished; the Task subscale focuses on the participant's agreement about the steps taken to help improve the client's situation; and the Bond subscale measures mutual liking and attachment by focusing on the tone of voice, empathy, and comfort in exploring intimate issues. The definitions of each level of the scales have been assembled into a 23-page coding manual (Darchuk et al., 2000). The intraclass correlation coefficient (ICC) of the Working Alliance Inventory – Observer Form assessed by two independent judges on 71 sessions was .63.

For the data analysis, we averaged the assessment made by the different raters on each of the instruments.

Procedure

The research design involved the following steps for each case: (1) formulation of the patient's plan; (2) identification of instances of the patient testing the therapist; (3) assessment of the clarity of testing; (4) assessment of the plan compatibility of the therapist's interventions, i.e. the degree to which the therapist's interventions passed or failed the patient's tests; (5) assessment of the patient's communications immediately before and after the intervention given by the therapist as a response to the test of the patient; and (6) assessment of the therapeutic alliance.

The raters

The study involved eleven independent judges who assessed the transcriptions using the instruments described in the previous section. The judges were divided into six groups and assigned to different instruments.

The judges comprised three males and eight females, with an average age of 35.5 years ($SD =$ 3.9; 29–42 years). Four of the aforementioned group had a PhD in dynamic and clinical psychology and had completed training in psychodynamic and/or integrated psychotherapy with 3 or more years of post-training clinical experience; five had a master's degree in clinical psychology and a specialization in psychodynamic psychotherapy with 3 or more years post-training clinical experience; and two had a master's degree in clinical psychology and were completing their training in psychodynamic or cognitive psychotherapy.

All judges completed a 98-hour course on the theoretical basis and clinical application of CMT, provided by the Control Mastery Theory- Italian Group (CMT-IG) before joining the present study. This course included 80 hours of training in the Plan Formulation Method and its application in clinical practice.

(1) Formulation of the Patient's Plan

To formulate the patient's plan, two independent judges followed the five steps of the Plan Formulation Method (Curtis & Silberschatz, 2022):

(1) The two judges reviewed the transcriptions of the first two sessions (the intake interview and first session) for each case and independently developed a formulation of each patient's plan. These formulations included the patient's goals for the therapy;the pathogenic beliefs impeding the attainment of goals that the patient wanted to disprove; the traumas from which these beliefs originated that the patient wanted to master; the possible ways the patient was likely to test the therapist; and the insights that might be helpful to the patient in disconfirming pathogenic beliefs. For each case, the judges created lists of goals, pathogenic beliefs, traumas, tests, and insights, and included not only items they believed were relevant to the patient but also any items they thought plausible but less relevant.

(2) The judges' lists were combined into master lists of goals, obstructions, traumas, tests, and insights for that case. The items were randomly distributed within the appropriate category.

(3) The master lists were returned to the judges, who independently rated the items with the Relevance Scale (Curtis et al., 1994) for their relevance to the patient;.

(4) The agreement between the judges' ratings was calculated for each patient's plan component using an intraclass correlation coefficient (ICC).

(5) For each case, the median rating for each category (i.e., goals, tests, etc.) was calculated. Items rated at or above the median were selected. Finally, another judge discarded any redundant items, and the remaining items were included in the formulation.

The patient's definitive plan formulation was used as basis to identify tests, rate the plan compatibility of the therapist's interventions, and evaluate segments of the patient's communications before and after the intervention given by the therapist as a response to the test of the patient.

(1) Identifying Tests

For each case, two judges independently read the verbatim transcriptions of each of the 71 sessions - except the intake interview and the first session - and identified all possible instances where the patient might be testing a pathogenic belief in the relationship with the therapist.

To select a pool of possible tests, judges were provided with a manual containing the definition of the test concept and of the testing strategies according to CMT, the criteria for identifying tests based on guidelines provided by Weiss (1993, p. 89), the instructions for identifying the limits of a patient testing communication, and the formulation of each patient's plan. Each testing sequence consisted of the patient's communications of a maximum length of 50 lines of text; their length varied according to the length of the test and the amount of information needed to understand the topic in the context of which the patient tested the therapist.

Each hour was read by each judge independently, thereby minimizing systematic bias and generating the maximum number of testing sequences. At the end of the process, a consensus between the two judges resulted in the selection of the testing sequences.

(1) Assessment of the Clarity of Testing

For each case, the testing sequences, along with the therapist's next intervention, were excerpted from the transcript, randomized, and presented to a group of three judges. These judges read the case formulation for each patient, a manual containing the definition of the test concept and of the testing strategies according to CMT, and the criteria for identifying tests based on guidelines provided by Weiss (1993, p. 89). Working independently, the judges rated each test using the Patient Scale of Testing (a revised version of the Patient Scale of Key Test; see Silberschatz et al., 1986). Only those segments to which all judges assigned a minimum score of 3 - indicating that the segment was a clear example of the patient testing the therapist - were included in the data analysis.

(1) Assessment of the Plan Compatibility of the Therapist's Interventions

The same three judges were also asked to independently rate the degree to which the therapist's response to each of the patient's tests either confirmed or disconfirmed the patient's pathogenic belief. The ratings, carried out using the Therapist Scale of Passing versus Failing Tests (Silberschatz et al., 1986), were based on the patient's plan formulation.

(1) Assessment of the Patient's Communications Immediately Before and After the Therapist's Intervention to the Patient's Test

The patient's immediate therapeutic progress was measured by rating segments of his/her verbalizations immediately before (pre-therapist's intervention, or baseline segment) and immediately after (post-therapist's intervention, or effect segment) each therapist's

intervention to the patient's test. These segments were extracted from the transcriptions without the interposed therapist's communications, randomized, and presented to two new judges. In this way, the judges were blind as to where the segment occurred in the session and whether the segment was a pre-therapist's intervention or post-therapist's intervention segment. The judges were also blind to treatment outcome. Each judge independently rated each segment using the following scales: the Overall Therapeutic Productivity scale of the APS (Waldron et al., 2004), the two DIS scales described above (Waldron et al., 2013); the Boldness Scale (Caston et al., 1986), and the Relaxation Scale (Curtis et al., 1986). These judges received the coding manual of each measure and the plan formulation of each patient on which to base their clinical judgements. They did not receive any specific training in the use of these measures.

(1) Assessment of the Therapeutic Alliance

For each case, two independent judges read the verbatim transcriptions of each session - except the intake interview and the first session-, totaling 71 sessions, and independently rated the therapeutic alliance using the Working Alliance Inventory – Observer Form (Darchuk et al., 2000). These judges received the WAI-O coding manual but did not receive any specific training in the use of this measure.

Statistical Analysis

A series of descriptive statistics were estimated for all variables included in the study (i.e., mean and standard deviation). In order to investigate the relationship among the variables assessing patient's communications before the therapist's interventions and among the variables assessing patient's communications after the therapist's interventions a correlation matrix was calculated. We then performed an exploratory factor analysis to see if these two sets of variables (pre-therapist's intervention patient's variables and post- therapist's intervention patient's variables) could be combined in two factors: patient's ease and productivity before therapist's intervention and patient's ease and productivity after therapist's intervention.

In order to examine the predictive power of therapists' ability to pass the patients' tests on patients' ease and therapeutic productivity, we applied a covariance design in which the dependent measure of the patients' ease and productivity after the therapists' interventions was predicted by the measure of the patients' ease and productivity before the therapists' interventions and the therapists' interventions evaluations. In this way, the effect of the therapists' interventions could be assessed while controlling for the effect of patients' ease and productivity before therapists' interventions. Finally, the correlation between the therapists' ability to pass the patients' tests and therapeutic alliance was calculated. Considering that our data set contained data nested within cases and therapists, a series of Mixed Models Analysis (MMA) were conducted.

The data analysis was performed using Jamovi (The Jamovi Project, Version 2.3.15).

Results

Descriptive Statistics

After reading the transcriptions of the therapy sessions for each case, two independent judges selected a pool of 707 incidents, i.e. instances in which they presumed that the patient might be testing a pathogenic belief in the relationship with the therapist. Of these, 171 tests were identified in the Millie case, 132 tests in the Hugo case, 184 tests in the Myra case, 114 tests in the Vicky case, and 106 tests in the Leon case.

Evaluations of these testing sequences by three independent judges with the Patient Scale of Testing led to the selection of a final sample of 339 tests. In order to narrow the focus of our study, we selected only those sequences of the patients' communications rated by all the judges as clear and/or excellent and clear-cut examples of the patient testing the therapist. . See Table I for the number of tests, therapist interventions, and pre- and post-therapist's interventions segments identified in each case.

Table II presents the descriptive statistics (mean and standard deviation) of all the variables studied in each case. All variables had distributions close to a normal curve (values of skewness and kurtosis were within ± 1).

Correlations and Factor Analysis

Due to the high level of correlations within the pre-therapist's intervention patient's variables ($r_{average}$ = .71) and the post-therapist's intervention patient's variables ($r_{average}$ = .82), we performed an exploratory factor analysis (principal component analysis) on the whole sample of patients' communications

Table I. Number of tests, therapist interventions, and pre- and post-therapist's interventions segments identified in each case.

	Millie	Hugo	Myra	Vicky	Leon	Total
Testing sequences	171	132	184	114	106	707
Tests (min. score ≥ 3)	110	48	77	54	50	339
Therapist's interventions	110	48	77	54	50	339
Pre- and post-therapist's interventions segments	220	96	154	108	100	678

in order to see if it was possible to synthesize the pre-therapist's intervention patient's variables in an overall "pre-therapist's intervention patient's ease and productivity" factor and the post-therapist's intervention patient's variables in an overall "post-therapist's intervention patient's ease and productivity" factor.

As we expected the factors to be moderately correlated, we chose a ProMax (oblique rotation). The KMO test of the sampling adequacy gave a result of .88 showing that the dimension of the sample was adequate. We followed Cattell's scree plot procedure (point of inflection of the curve) and the criterion of eigenvalues > 1 for selecting the factors. We obtained a two-factor solution explaining 82.2% of the common variance. The first unrotated factor had an eigenvalue of 5.28 and explained 52.80% of the common variance; the second factor had an eigenvalue of 2.94 and explained 29.43% of the common variance. Table III presents the Promax-rotated principal component analysis factor pattern loadings. The correlations between the two factors was r = .28. Based on the factor pattern loadings, Factor 1 can be labeled as *Patient's ease and productivity before therapist's intervention*; Factor 2 as *Patient's ease and productivity after therapist's intervention*.

Mixed Model Analyses

In order to test the predictive value of therapists' ability to pass the patients' tests on patients' ease and therapeutic productivity, a Mixed Model Analysis was conducted in which case intercept and therapist intercept served as random effects. Our data were in fact nested within the cases ($n = 5$) and within the therapists ($n = 4$) (Level 2 variables). We first conducted this analysis on the entire sample of 339 observations (Level 1 variables). The independent variable were the patient's ease and productivity before the therapist's intervention and the therapist's intervention evaluation. The dependent variable was the patient's ease and productivity after therapist's intervention. The fixed effects of these analyses are presented in Table IV. As can be seen, the therapists' response evaluation significantly predicted patient's ease and productivity after therapist's intervention after controlling for the effect of patient's ease and productivity before therapist's intervention.

To test our second hypothesis, a MMA was conducted in which case intercept and therapist intercept served as random effects. We applied the same covariance design but on the mean score per session of the patient's and therapist's variables ($N = 67$).[2] The independent variable was the mean score per-session of patient's ease and productivity factor before therapist's intervention and the mean score per-session of therapist's intervention evaluation. The dependent variable was the mean score per-session of the patient's ease and productivity factor after therapist's intervention. The fixed effects of these analyses are presented in Table V. As can be seen, the therapist response evaluation in a session significantly predicted patient's ease and productivity after therapist's intervention within the same sessions after controlling for the effect of patient's ease and productivity before therapist's intervention.

To test our third hypothesis, a MMA was conducted in which both a case intercept and therapist intercept served as random effects, the mean score per session of therapist's interventions as independent variables, and the mean score per session of the therapeutic alliance as dependent variable ($N = 67$). The fixed effects are presented in Table VI. As can be seen, the therapist response evaluation in a session significantly correlated with the therapeutic alliance during the same session.

To sum up, all our hypotheses we confirmed.

Discussion

Throughout the testing activity, patients enacted in their relationship with the therapist an emotional conflict originally experienced in relation to their significant others. It pitted a healthy, adaptive goal against the fear that pursuing it would hurt significant others, would end up in the patient being hurt, or would threaten their relationship. Patients test their therapist to assess the level of safety of the therapeutic relationship in the hope that the therapist would support them in the pursuit of their goals,

Table II. Descriptive statistics.

	Millie		Hugo		Myra		Vicky		Leon	
	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*	Mean	*SD*
Therapist interventions	3.79	1.16	2.76	.67	3,78	.79	3.55	.73	3.05	.62
DIS_7 Pre	2.25	.58	1.72	.54	1,82	.46	1.95	.48	2.18	.53
DIS_7 Post	1.90	.88	1.19	.82	1,51	.58	1.52	.65	1.72	.70
DIS_9 Pre	1.77	.73	1.49	.50	1,81	.51	1.92	.47	2.21	.52
DIS_9 Post	1.45	.86	.99	.70	1,51	.67	1.48	.69	1.75	.74
APS Pre	1.87	.66	1.29	.58	1,55	.57	1.62	.45	2.00	.58
APS Post	1.58	.83	.88	.66	1,31	.61	1.34	.69	1.62	.71
Bold Pre	5.62	1.15	4.60	.87	5,47	.90	5.70	.90	5.84	1.02
Bold Post	5.03	1.59	3.59	1.69	4,71	1.23	4.81	1.07	5.02	1.34
Relaxation Pre	2.35	.67	1.90	.47	2,42	.41	2.33	.45	2.66	.55
Relaxation Post	2.00	.90	1.38	.89	2,01	.75	1.94	.75	2.18	.75

Note. Millie $n = 110$; Hugo $n = 48$; Myra $n = 77$; Vicky $n = 54$; Leon $n = 50$.

disconfirming the pathogenic beliefs that hinder such a pursuit and thereby offering them a corrective emotional experience (Castonguay & Hill, 2012).

In keeping with previous empirical research (Horowitz et al., 1975; Silberschatz, 1986; Silberschatz et al., 1986; Silberschatz & Curtis, 1993) the results of the present study indicate that when the therapists' interventions passed a patients' tests, the patients showed signs of immediate improvement, feeling safer and becoming more therapeutically productive (i.e., more able to oscillate between experiencing and reflecting on their own experiences; more able to deepen their exploration of significant personal issues with boldness, ease, and interest; more able to address their problematic patterns of relating and feelings in a productive manner; and more involved in their relationship with the therapist).

Table III. Promax-rotated principal component analysis factor pattern loadings.

	Factors	
	Patient's ease and productivity before therapist's intervention	Patient's ease and productivity after therapist's intervention
DIS_7 Post		0.960
DIS_9 Post		0.930
APS Post		0.933
Bold Post		0.880
Relaxation Post		0.959
DIS_7 Pre	0.891	
DIS_9 Pre	0.868	
APS Pre	0.908	
Bold Pre	0.843	
Relaxation Pre	0.875	

Note. $N = 339$. Loadings are sorted by magnitude. Loadings Below |.30| are suppressed for clarity. All displayed loadings were significant at I < .05.

As was the case in Silberschatz and Curtis (1993) study, the correlations were stronger when the average level of each variable in each session was taken into consideration. This might be explained by the presence of the *cumulative* effect of the therapist's ability to disconfirm the patient's pathogenic beliefs within a session.

Finally, the results show that the average ability of the therapists to pass their patients' tests in each session is correlated with the therapeutic alliance, so it may be seen as being influenced by it, or at least one of the contributing factors of the alliance.

Tests require specific responses from the therapist to be passed, responses that are tailored to the patient's particular needs; for this reason, therapists' ability to pass the tests of their patients can be seen as an indicator of their responsiveness to patients (Silberschatz, 2021). Moreover, to detect and understand patients' tests it is important to know their goals, pathogenic beliefs, traumas, and testing strategies. Having accurate plan formulation, made possible thanks to methods such as the Plan Formulation Method, allows the therapist to predict the ways the patient may test them and how to respond. In other words, it helps therapists to design and provide case-specific treatments, to develop a stronger alliance with their patients and to be more accurately responsive (Critchfield et al., 2022).

Limitations and Future Research

The present study, even if it is the largest single research study on testing so far conducted, has several limitations. First, the level of reliability of the assessments conducted on five rating scales was less than .70 (i.e., .60, .68, .63, .59, and .63),

Table IV. Relationship between the therapist's interventions evaluations and patient's ease and productivity after therapist's interventions (N = 339).

	MMA of dependent variables					
	Estimate	*SE*	*df*	*t*	*p*	*r*
Patient's ease and productivity before therapist's intervention	0.11	0.08	315.17	4.08	<.001	0.22
Therapist Intervention	0.18	0.48	310	3.68	<.001	**0.20**

Note. Dependent Variable: patient's ease and productivity after therapist's intervention.

which was perhaps because we were not able to provide the raters with specific training on each measure. In fact, previous studies using the same scales whose reliability was lower in this study were able, with appropriate training provided to the raters, to obtain higher levels of reliability (see, for example, Waldron et al., 2013; Gazzillo et al., 2014; Gazzillo et al., 2018).

Secondly, the correlations between the therapists' ability to pass the patients' tests and the ease and productivity of the following patients' communications —when assessed on a single-instance basis—was low (average $r = .20$). However, if we consider the correlations between the average therapists' ability to pass their patients' tests in each session, the correlation was stronger (average $r = .36$). This, in our opinion, suggests a cumulative effect of the therapists' ability to pass their patients tests.

Another limitation of this study derives from the fact that it was based only in the analysis of the content of transcribed session, i.e. on the verbal behavior of patients and therapists. In a study with some similarity, using a different but in many ways similar approach of assessing responsiveness to the present study, the nonverbal therapist behavior has been found to have a stronger correlation with the outcome than the verbal content of what the therapist says (Caspar et al., 2005). Moreover, in an attempt to disentangle process and outcome with the original material of the Silberschatz et al. (1986) study, even with the limitations inherent to transcript material, Caspar et al. (2000) assessed the plan compatibility related to process (how the therapist intervenes) as opposed to the plan compatibility related to content (what he/she says) , and the results suggest that the process is more important in some, and the content in other cases.

Finally, the study did not allow us to confirm the direction of influence of the therapists' ability to pass the patients' tests and the therapeutic alliance. It is therefore not clear what influenced what or whether there was a circular relationship between the two constructs.

Future studies involving larger samples of patients with problems of different levels of severity and treated by therapists with different theoretical orientations might support CMT hypotheses on the effects of the therapist's ability to pass their patient's tests and the correlation between this and the outcome of the therapy. Indirect evidence of a correlation was provided by Silberschatz (2017), who studied the whole sample of patients in the Mount Zion Brief Psychotherapy Project. They concluded that therapists' ability to provide plan-compatible interventions that disconfirmed patients' pathogenic beliefs (of which the ability of the therapist to pass the patient's test is one instance) correlated with the outcomes of their therapies. However, studies comparing the difference between the therapist's ability to pass their patient tests in a larger sample of good- and poor-outcome treatments would be

Table V. Relationship between the therapist's interventions evaluations and patient's ease and productivity after therapist's interventions (N = 67).

	MMA of dependent variables					
	Estimate	*SE*	*df*	*t*	*p*	*r*
Patient's ease and productivity before therapist's intervention	0.45	0.16	64	3.26	0.005	0.37
Therapist Intervention	0.34	0.11	64	3.13	0.003	**0.36**

Note: Dependent Variable: Patient's ease and productivity after therapist's intervention.

Table VI. Relationship between the therapist's interventions evaluations and the therapeutic alliance (N = 67).

	MMA of dependent variables					
	Estimate	*SE*	*df*	*t*	***p***	***r***
Dependent variable: Therapeutic alliance						
Therapist interventions	13.9	6.56	64	2.12	.038	**0.25**

useful for understanding if passing the test is a core element of an effective therapeutic work.

Future studies using a generalized method of moments for dynamic panel data (Falkenström et al., 2016) might also assess the direction of the relationship between the ability of the therapist to pass their patient's tests and the therapeutic alliance.

Acknowledgements

The authors express their sincere gratitude to the clinicians and researchers who generously volunteered many hours to complete of the clinical ratings. This study would not have been possible without their help and expertise.

Funding

The study was supported by Sapienza University of Rome grant number AR120172AD5F50F4.

Notes

[1] https://www.sfprg.org

[2] In four session of the overall sample (N = 71), the judges were not able to identify good and/or excellent, clear-cut examples of the patient testing the therapist.

Disclosure Statement

No potential conflict of interest was reported by the author(s).

Ethical Approval

The study was conducted in accordance with the 1964 Helsinki declaration and was approved by the Local Ethics Committee of the Department of Dynamic and Clinical Psychology and Health Studies, Sapienza University of Rome (protocol number: 0000832).

Availability of Data and Material

The data that support the findings of this study are available from the corresponding author (FG), upon reasonable request.

Author Contributions

Ramona Fimiani and Francesco Gazzillo contributed to the conceptualization and design of the study and agreed to ensure that the questions that were related to the accuracy or integrity of any part of the work were appropriately investigated and resolved. They are primarily accountable for all aspects of the work. Bernard Gorman analyzed the data and oversaw the composition of the statistical parts of the manuscript. Francesco Gazzillo revised the paper for intellectual content and approved its final version. Jessica Leonardi, Giuseppe Stefano Biuso, Martina Rodomonti, Camilla Mannocchi, and Federica Genova contributed to data acquisition and it's subsequent interpretation.

ORCID

FRANCESCO GAZZILLO http://orcid.org/0000-0002-7955-3710

JESSICA LEONARDI http://orcid.org/0000-0002-9305-0554

References

Aafjes-van Doorn, K., Kamsteeg, C., & Silberschatz, G. (2020). Cognitive mediators of the relationship between adverse childhood experiences and adult psychopathology: A systematic review. *Development and Psychopathology*, *32*(3), 1017–1029. https://doi.org/10.1017/S0954579419001317

Ahn, H., & Wampold, B. E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology*, *48*(3), 251–257. https://doi.org/10.1037/0022-0167.48.3.251

Alexander, F., & French, T. M. (1946). *Psychoanalytic therapy; principles and application*. Ronald Press.

American Psychological Association. (2012). Recognition of psychotherapy effectiveness. https://www.apa.org/about/policy/resolution-psychotherapy.

Bacal, H. A. (1998). Optimal responsiveness and the therapeutic process. In H. A. Bacal (Ed.), *Optimal responsiveness: How therapists heal their patients* (pp. 3–34). Jason Aronson.

Barber, J. P., & Solomonov, N. (2019). Toward a personalized approach to psychotherapy outcome and the study of therapeutic change. *World Psychiatry*, *18*(3), 291–292. https://doi.org/10.1002/wps.20666

Barkham, M., & Lambert, M. J. (2021). The efficacy and effectiveness of psychological therapies. In M. Barkham, W. Lutz, & L. G. Castonguay (Eds.), *Bergin and Garfield's handbook of psychotherapy and behavior change: 50th anniversary edition* (pp. 135–189). John Wiley & Sons, Inc.

Bugas, J., McCollum, J., Kealy, D., Silberschatz, G., Curtis, J. T., & Reid, J. (2021). Identifying patient verbal coaching in psychotherapy. *Counselling and Psychotherapy Research*, 1–11. https://doi.org/10.1002/capr.12495

Bugas, J., & Silberschatz, G. (2000). How patients coach their therapists in psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, *37*(1), 64–70. https://doi.org/10.1037/h0087676

Caspar, F., Grossmann, C., Unmüssig, C., & Schramm, E. (2005). Complementary therapeutic relationship: Therapist behavior, interpersonal patterns, and therapeutic effects. *Psychotherapy Research*, *15*(1-2), 91–102. https://doi.org/10.1080/10503300512331327074

Caspar, F., Pessier, J., Stuart, J., Safran, J. D., Samstag, L. W., & Guirguis, M. (2000). One step further in assessing how interpretations influence the process of psychotherapy. *Psychotherapy Research*, *10*(3), 309–320. https://doi.org/10.1080/713663734

Caston, J., Goldman, R. K., & McClure, M. M. (1986). The immediate effects of psychoanalytic interventions. In J. Weiss, & H. Sampson & the Mount Zion Psychotherapy Research Group (Eds.), *The psychoanalytic process: Theory, clinical observation, and empirical research* (pp. 277–298). Guilford Press.

Castonguay, L. G., Constantino, M. J., & Beutler, L. E. (2019). *Principles of change: How psychotherapists implement research in practice*. Oxford University Press.

Castonguay, L. G., & Hill, C. E.2012). *Transformation in psychotherapy: Corrective experiences across cognitive behavioral, humanistic, and psychodynamic approaches*. American Psychological Association. https://doi.org/10.1037/13747-000

Critchfield, K. L., Gazzillo, F., & Kramer, U. (2022). Case formulation of interpersonal patterns and its impact on the therapeutic process: Introduction to the issue. *Journal of Clinical Psychology*, *78*(3), 379–385. https://doi.org/10.1002/jclp.23322

Cuijpers, P., Quero, S., Noma, H., Ciharova, M., Miguel, C., Karyotaki, E., Cipriani, A., Cristea, I. A., & Furukawa, T. A. (2021). Psychotherapies for depression: a network meta-analysis covering efficacy, acceptability and long-term outcomes of all main treatment types. *World Psychiatry*, *20*(2), 283–293. https://doi.org/10.1002/wps.20860

Curtis, J. T., Ransohoff, P., Sampson, F., Brumer, S., & Brontstein, A. A. (1986). Expressing warded-off contents in behavior. In J. Weiss, H. Sampson & Mount Zion Psychotherapy Research Group (Eds.), *The psychoanalytic process: Theory, clinical observation, and empirical research* (pp. 187–205). Guilford Press.

Curtis, J. T., & Silberschatz, G. (2022). The plan formulation method. In T. D. Eells (Ed.), *Handbook of psychotherapy case formulation* (3rd ed., pp. 88–112). Guilford Press.

Curtis, J. T., Silberschatz, G., Sampson, H., & Weiss, J. (1994). The plan formulation method. *Psychotherapy Research*, *4*(3-4), 197–207. https://doi.org/10.1080/10503309412331334032

Darchuk, A., Wang, V., Weibel, D., Fende, J., Anderson, T., & Horvath, A. O. (2000). Manual for the Working Alliance Inventory—Observer Form, 4th revision. Unpublished Manuscript.

Dragioti, E., Karathanos, V., Gerdle, B., & Evangelou, E. (2017). Does psychotherapy work? An umbrella review of meta-analyses of randomized controlled trials. *Acta Psychiatrica Scandinavica*, *136*(3), 236–246. https://doi.org/10.1111/acps.12713

Falkenström, F., Ekeblad, A., & Holmqvist, R. (2016). Improvement of the working alliance in one treatment session predicts improvement of depressive symptoms by the next session. *Journal of Consulting and Clinical Psychology*, *84*(8), 738–751. https://doi.org/10.1037/ccp0000119

Fimiani, R., Gazzillo, F., Fiorenza, E., Rodomonti, M., & Silberschatz, G. (2020). Traumas and their consequences according to control-mastery theory. *Psychodynamic Psychiatry*, *48*(2), 113–139. https://doi.org/10.1521/pdps.2020.48.2.113

Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy*, *55*(4), 316–340. https://doi.org/10.1037/pst0000172

Frank, J. D., & Frank, J. B. (1991). *Persuasion and healing: A comparative study of psychotherapy* (3rd ed.). Johns Hopkins University Press.

Gazzillo, F. (2021). *Fidarsi dei pazienti. Introduzione alla Control-Mastery Theory [Trusting patients. Introduction to control-mastery theory]*. Raffaello Cortina.

Gazzillo, F. (2022). Toward a more comprehensive understanding of pathogenic beliefs: Theory and clinical implications. *Journal of Contemporary Psychotherapy*, https://doi.org/10.1007/s10879-022-09564-5

Gazzillo, F., Curtis, J., & Silberschatz, G. (2022). The plan formulation method: An empirically validated and clinically useful procedure applied to a clinical case of a patient with a severe personality disorder. *Journal of Clinical Psychology*, *78*(3), 409–421. https://doi.org/10.1002/jclp.23299

Gazzillo, F., Dimaggio, G., & Curtis, J. T. (2021). Case formulation and treatment planning: How to take care of relationship and symptoms together. *Journal of Psychotherapy Integration*, *31*(2), 115–128. https://doi.org/10.1037/int0000185

Gazzillo, F., Fimiani, R., De Luca, E., Dazzi, N., Curtis, J. T., & Bush, M. (2020). New developments in understanding morality: Between evolutionary psychology, developmental psychology, and control-mastery theory. *Psychoanalytic Psychology*, *37*(1), 37–49. https://doi.org/10.1037/pap0000235

Gazzillo, F., Genova, F., Fedeli, F., Curtis, J. T., Silberschatz, G., Bush, M., & Dazzi, N. (2019). Patients' unconscious testing activity in psychotherapy: A theoretical and empirical overview. *Psychoanalytic Psychology*, *36*(2), 173–183. https://doi.org/10.1037/pap0000227

Gazzillo, F., Waldron, S., Genova, F., Angeloni, F., Ristucci, C., & Lingiardi, V. (2014). An empirical investigation of analytic process: Contrasting a good and poor outcome case. *Psychotherapy*, *51*(2), 270–282. https://doi.org/10.1037/a0035243

Gazzillo, F., Waldron, S., Gorman, B. S., Stukenberg, K., Genova, F., Ristucci, C., Faccini, F., & Mazza, C. (2018). The components of psychoanalysis: Factor analyses of process measures of 27 fully recorded psychoanalyses. *Psychoanalytic Psychology*, *35*(2), 184–195. https://doi.org/10.1037/pap0000155

Horowitz, L. M., Sampson, H., Siegelman, E. Y., Wolfson, A., & Weiss, J. (1975). On the identification of warded-off mental contents: An empirical and methodological contribution. *Journal of Abnormal Psychology*, *84*(5), 545–558. https://doi.org/10.1037/h0077139

Kealy, D., McCollum, J., Curtis, J. T., Silberschatz, G., Katie, A., & Luo, X. (2022). Failure to respond to the patient's coaching: A case study of premature termination in psychodynamic psychotherapy. *Counselling Psychology Quarterly*, *35*(4), 789–813. https://doi.org/10.1080/09515070.2021.2000941

Kramer, U.2019). *Case formulation for personality disorders: Tailoring psychotherapy to the individual client*. Elsevier Academic Press.

Kramer, U. (2020). Individualizing psychotherapy research designs. *Journal of Psychotherapy Integration*, *30*(3), 440–457. https://doi.org/10.1037/int0000160

Kramer, U. (2021). Therapist responsiveness in treatments for personality disorders. In J. C. Watson, & H. Wiseman (Eds.), *The responsive psychotherapist: Attuning to clients in the moment* (pp. 237–255). American Psychological Association. https://doi.org/10.1037/0000240-012

Kramer, U., & Stiles, W. B. (2015). The responsiveness problem in psychotherapy: A review of proposed solutions. *Clinical Psychology: Science and Practice*, *22*(3), 277–295. https://doi.org/10.1111/cpsp.12107

Lambert, M. J., & Barley, D. E. (2001). Research summary on the therapeutic relationship and psychotherapy outcome. *Psychotherapy: Theory, Research, Practice, Training*, *38*(4), 357–361. https://doi.org/10.1037/0033-3204.38.4.357

Laska, K. M., Gurman, A. S., & Wampold, B. E. (2014). Expanding the lens of evidence-based practice in psychotherapy: A common factors perspective. *Psychotherapy*, *51*(4), 467–481. https://doi.org/10.1037/a0034332

Leonardi, J., Gazzillo, F., & Dazzi, N. (2022). The adaptive unconscious in psychoanalysis. *International Forum of Psychoanalysis*, *31* (4), 201–217. https://doi.org/10.1080/0803706X.2021.1893382

Norcross, J. C., & Lambert, M. J.2019). *Psychotherapy relationships that work: Vol. 1. Evidence-based therapist contributions* (3rd ed.). Oxford University Press.

Norcross, J. C., & Wampold, B. E.2019). *Psychotherapy relationships that work: Vol. 2. Evidence-based therapist responsiveness* (3rd ed.). Oxford University Press.

Novak, A. N., Luedemann, J., & Andreas, S. (2022). When patients probe the analyst: Manifestations of patient testing and its complexity – An in-depth exploration of case examples of extant research. *International Forum of Psychoanalysis*, 1–17. https://doi.org/10.1080/0803706X.2022.2075564

Scharf, R. D., Waldron, S., Firestein, S. K., Goldberger, M., & Burton, A. (2010). The Analytic Process Scales (APS) coding manual. Unpublished manuscript. (Original work 1999).

Shedler, J. (2018). Where Is the evidence for "evidence-based" therapy? *Psychiatric Clinics of North America*, *41*(2), 319–329. https://doi.org/10.1016/j.psc.2018.02.001

Silberschatz, G. (2017). Improving the yield of psychotherapy research. *Psychotherapy Research*, *27*(1), 1–13. https://doi.org/10.1080/10503307.2015.1076202

Silberschatz, G. (2021). Responsiveness in control-mastery theory. In J. C. Watson, & H. Wiseman (Eds.), *The responsive psychotherapist: Attuning to clients in the moment* (pp. 133–150). American Psychological Association. https://doi.org/10.1037/0000240-007

Silberschatz, G., & Aafjes-van Doorn, K. (2017). Pathogenic beliefs mediate the relationship between perceived negative parenting and psychopathology symptoms. *Journal of Aggression, Maltreatment & Trauma*, *26*(3), 258–275. https://doi.org/10.1080/10926771.2016.1259279

Silberschatz, G., & Curtis, J. T. (1986). Clinical implications of research on brief dynamic psychotherapy II. How the therapist helps or hinders therapeutic progress. *Psychoanalytic Psychology*, *3*(1), 27–37. https://doi.org/10.1037/0736-9735.3.1.27

Silberschatz, G., & Curtis, J. T. (1993). Measuring the therapist's impact on the patient's therapeutic progress. *Journal of Consulting and Clinical Psychology*, *61*(3), 403–411. https://doi.org/10.1037/0022-006X.61.3.403

Silberschatz, G., Curtis, J. T., Sampson, H., & Weiss, J. (1991). Mount zion hospital and medical center: Research on the process of change in psychotherapy. In L. E. Beutler, & M. Crago (Eds.), *Psychotherapy research: An international review of programmatic studies* (pp. 56–64). American Psychological Association. https://doi.org/10.1037/10092-006

Silberschatz, G., Sampson, H., & Weiss, J. (1986). Testing pathogenic beliefs versus seeking transference gratifications. In J. Weiss, H. Sampson & the Mount Zion Psychotherapy Research Group (Eds.), *The psychoanalytic process: Theory, clinical observation and empirical research* (pp. 267–276). Guilford Press.

Silberschatz, G. (1986). Testing pathogenic beliefs. In J. Weiss, H. Sampson & the Mount Zion Psychotherapy Research Group (Eds.), *The psychoanalytic process: Theory, clinical observation and empirical research* (pp. 256–266). Guilford Press.

Silberschatz, G. (2005). *Transformative relationships: The control-mastery theory of psychotherapy*. Routledge.

Steinert, C., Munder, T., Rabung, S., Hoyer, J., & Leichsenring, F. (2017). Psychodynamic therapy: As efficacious as other empirically supported treatments? A meta-analysis testing equivalence of outcomes. *American Journal of Psychiatry*, *174*(10), 943–953. https://doi.org/10.1176/appi.ajp.2017.17010057

Stiles, W. B. (2021). Responsiveness in psychotherapy research: Problems and ways forward. In J. C. Watson, & H. Wiseman (Eds.), *The responsive psychotherapist: Attuning to clients in the moment* (pp. 15–35). American Psychological Association. https://doi.org/10.1037/0000240-002

Stiles, W. B., Honos-Webb, L., & Surko, M. (1998). Responsiveness in psychotherapy. *Clinical Psychology: Science and Practice*, *5*(4), 439–458. https://doi.org/10.1111/j.1468-2850.1998.tb00166.x

Waldron, S. (2011). Dynamic Interaction Scales (DIS) Manual from the APS Group. Unpublished manuscript.

Waldron, S., Gazzillo, F., Genova, F., & Lingiardi, V. (2013). Relational and classical elements in psychoanalyses: An empirical study with case illustrations. *Psychoanalytic Psychology*, *30* (4), 567–600. https://doi.org/10.1037/a0033959

Waldron, S., Scharf, R. D., Hurst, D., Firestein, S. K., & Burton, A. (2004). What happens in a psychoanalysis? *The International Journal of Psychoanalysis*, *85*(2), 443–466. https://doi.org/10.1516/5PPV-Q9WL-JKA9-DRCK

Wampold, B. E. (2015). How important are the common factors in psychotherapy? An update. *World Psychiatry*, *14*(3), 270–277. https://doi.org/10.1002/wps.20238

Wampold, B. E., & Imel, Z. E. (2015). *The great psychotherapy debate: The evidence for what makes psychotherapy work* (2nd ed.). Routledge/Taylor & Francis Group.

Wampold, B. E., Mondin, G. W., Moody, M., Stich, F., Benson, K., & Ahn, H.-n. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empiricially, "all must have prizes. *Psychological Bulletin*, *122*(3), 203–215. https://doi.org/10.1037/0033-2909.122.3.203

Weiss, J. (1990). The centrality of adaptation. *Contemporary Psychoanalysis*, *26*(4), 660–676. https://doi.org/10.1080/00107530.1990.10746684

Weiss, J. (1993). *How psychotherapy works*. Guilford Press.

Weiss, J. (1994). The analyst's task. *Contemporary Psychoanalysis*, *30*(2), 236–254. https://doi.org/10.1080/00107530.1994.10746853

Weiss, J. (1997). The role of pathogenic beliefs in psychic reality. *Psychoanalytic Psychology*, *14*(3), 427–434. https://doi.org/10.1037/h0079734

Weiss, J., Sampson, H., & Mount Zion Psychotherapy Research Group. (1986). *The psychoanalytic process: Theory, clinical observation, and empirical research*. Guilford Press.

Westen, D., Novotny, C. M., & Thompson-Brenner, H. (2004). The empirical status of empirically supported psychotherapies: Assumptions, findings, and reporting in controlled clinical trials. *Psychological Bulletin*, *130*(4), 631–663. https://doi.org/10.1037/0033-2909.130.4.631

Wu, M. B., & Levitt, H. M. (2020). A qualitative meta-analytic review of the therapist responsiveness literature: Guidelines for practice and training. *Journal of Contemporary Psychotherapy*, *50*(3), 161–175. https://doi.org/10.1007/s10879-020-09450-y

Xu, H., & Tracey, T. J. G. (2015). Reciprocal influence model of working alliance and therapeutic outcome over individual therapy course. *Journal of Counseling Psychology*, *62*(3), 351–359. https://doi.org/10.1037/cou0000089

Zilcha-Mano, S. (2017). Is the alliance really therapeutic? Revisiting this question in light of recent methodological advances. *American Psychologist*, *72*(4), 311–325. https://doi.org/10.1037/a0040435

Zilcha-Mano, S. (2021). Toward personalized psychotherapy: The importance of the trait-like/state-like distinction for understanding therapeutic change. *American Psychologist*, *76*(3), 516–528. https://doi.org/10.1037/amp0000629