

The assessment of therapist responsiveness in psychotherapy research: a systematic review

Giovanna Esposito, Francesco Cuomo, Angela Di Maro, Raffaella Passeggia

Department of Humanities, University of Naples Federico II, Italy

Correspondence: Raffaella Passeggia, Department of Humanities, University of Naples Federico II, Via Porta di Massa, 1 – 80133 Naples, Italy.
E-mail: raffaella.passeggia@unina.it

Contributions: GE, is responsible for the conceptualization of the study and provided supervision in all the steps of the study and manuscript writing; GE, FC, RP, proceeded with steps for the analysis and contributed to the interpretation of results; FC, ADM, wrote the first draft of the manuscript. All the authors contributed to the subsequent redrafting and critical review of the manuscript, contributed to and approved the final manuscript.

Conflict of interest: all authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Ethics approval and consent to participate: not applicable.

Funding: none.

Availability of data and materials: the data reported in this manuscript have been presented only in this study and will be made available under reasonable request.

Citation: Esposito, G., Cuomo, F., Di Maro, A. & Passeggia, R., (2024). The assessment of therapist responsiveness in psychotherapy research: a systematic review. *Research in Psychotherapy: Psychopathology, Process and Outcome*, 27(2), 751. doi: 10.4081/ripppo.2024.751

Received: 16 January 2024.
Accepted: 13 May 2024.

Publisher's note: all claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article or claim that may be made by its manufacturer is not guaranteed or endorsed by the publisher.

©Copyright: the Author(s), 2024
Licensee PAGEPress, Italy
Research in Psychotherapy: Psychopathology, Process and Outcome 2024; 27:751
doi:10.4081/ripppo.2024.751

This article is distributed under the terms of the Creative Commons Attribution-NonCommercial International License (CC BY-NC 4.0) which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

ABSTRACT

Therapist responsiveness is an emerging construct in psychotherapy research that still lacks a clear definition and, consequently, a unique operationalization. Indeed, there is a great overlap between therapist responsiveness and other variables, such as attunement, flexibility, and empathy. This overlap inevitably hinders the assessment of the therapist's responsiveness, although it is crucial for the effectiveness of the treatments. Therefore, the current systematic review aims at exploring the different methodologies of measurement and analysis of therapist responsiveness, including both direct and indirect instruments. The results confirmed difficulties and divergences in the operationalization of the construct, as shown by the great heterogeneity found in the choice and use of the tools. Furthermore, this review provides guidance for future research, clinical practice, and training of therapists.

Key words: therapist responsiveness, evaluation, measurement instruments, effectiveness, efficacy.

Introduction

Therapist responsiveness (TR) is an emerging construct in psychotherapy research, and consequently, there is still no agreement on its definition and operationalization. One of the first and most shared formulations is the one by Stiles *et al.* (1998) which conceptualized it as a therapist's interpersonal competence, defining it as the set of behaviors influenced by the emerging therapeutic context. Subsequently, Elkin *et al.* (2014) have extended this definition by indicating TR as the degree to which the therapist shows himself attentive, caring, and respectful towards the patient, keeping in mind his uniqueness; he attempts to recognize and understand the patient's emerging concerns; and he appears clearly interested in the patient's communications, both in terms of content and feelings. A following conceptualization is that by Hatcher (2015), according to which TR refers to the therapist's ability to achieve benefits for a patient by adjusting the interventions to the current state of the patient and the therapeutic process. In an attempt to combine all these definitions, Watson and Wiseman (2021) have recently conceptualized TR as a therapist's interpersonal competence consisting of a set of flexible interventions aligned with patients' emotional and cognitive states and being able to respond to several contingencies within a specific clinical relationship.

Indeed, framing TR is a difficult task due to its similarity with other constructs, such as attunement, flexibility, and empathy,

among others; in fact, all of them require a connection to the patient's mental state, but there are significant differences. In particular, the term "attunement" refers to a tuning process that allows therapists to genuinely connect with patients' feelings and experiences (Erskine, 1998) without necessarily responding to them, whereas TR is not a process but a competence, and in addition, it concerns what the therapist does during the session. Therefore, the attunement could be considered a necessary but not sufficient precondition for TR: a therapist who is responsive is also tuned to the patient, but a tuned therapist is not necessarily responsive. Flexibility refers to a therapist's personal characteristics that allows them to adapt the treatment to patients' needs and individuality (Cahill *et al.*, 2009); therefore, training therapists to flexibility is more difficult, in contrast to TR, which is an interpersonal competence that can be easier taught. Finally, empathy refers to the ability to perceive, recognize, and identify oneself with other people's emotions. Undoubtedly, empathy is an important prerequisite for TR, but even empathizing does not necessarily imply responding to the other's emotional state. By contrast, as claimed by Stiles *et al.* (1998), TR consists of a general and positive trend within the clinical relationship to act to produce several desired results; in this regard, they use the term "appropriate responsiveness", which is equivalent to "do the right thing in the right moment" (Stiles & Horvath, 2017). Thus, it is possible to emphasize that the aspect that, to date, better identifies TR and distinguishes it from other constructs is this behavioral dimension, which consists of the therapist's actions (*e.g.*, what the therapist does) in response to the patient's mental states. These therapeutic responses or interventions may be verbal, non-verbal, and paraverbal, such as empathic listening attitudes or interpretations, and may simultaneously deal with different interaction levels within the clinical relationship (*e.g.*, tone of voice, the patient's regressive movements, his or her emerging emotions, the presence of new symptoms, what happened in previous sessions, and cultural, economic, ethnic, or gender differences) (Stiles *et al.*, 1998). Furthermore, according to these authors, TR occurs within multiple time intervals, on time scales ranging from months to milliseconds.

According to some authors, TR is a problematic concept for a variety of reasons. Firstly, it leads therapists to implement different treatments despite guidelines provided by theoretical manuals, characterizing itself as opposed to one of the dimensions of treatment integrity, namely adherence (Stiles, 2021). However, several researchers have shown that therapists can adhere to a treatment protocol and at the same time adapt their interventions to patients' needs (Kendall & Frank, 2018; Marques *et al.*, 2019) with a positive impact on the therapeutic outcome (Esposito *et al.*, 2024; Esposito *et al.*, 2020), supporting the idea that manualized treatments may incorporate methods and techniques of other models (McAleavey & Castonguay, 2014). In this regard, Hatcher (2021) underlines the importance of integrating these two factors in order to tailor the treatment to patients' necessities while remaining within a clear theoretical framework. Secondly, TR is highly affected by participants' subjective characteristics; thus, for instance, therapists could think they were responsive because they implemented specific interventions (*i.e.*, interpretation) in line with what was expected from their own theoretical model, but a patient could judge therapists' behavior as not responsive, since it was painful (Kivlighan *et al.*, 2017).

Therefore, several studies showed that the decisive aspect of treatment effectiveness is the therapeutic relationship. In this perspective, Reis (2014) identifies TR as a key element for establishing and maintaining therapeutic relationships, since it consists of

interacting with patients to facilitate achieving their objectives. For this reason, TR may play an important role in fostering working alliance (WA), especially in identifying and repairing alliance ruptures (Eubanks *et al.*, 2021). This is extremely important considering that several studies showed an association between the reparative processes of alliance ruptures and positive outcomes, *e.g.*, clients' symptomatologic improvements and a significant reduction in early drop-out (Eubanks *et al.*, 2019). Indeed, researchers also found that TR may have a direct and positive impact on treatment effectiveness (Zuroff *et al.*, 2010).

In light of what was discussed, it is clear that enhancing and developing TR is essential for psychotherapists. In particular, it is also necessary to develop a set of valid instruments that allow for the measurement of TR, which to date is still lacking in the psychotherapy literature. As noted by Stiles (2021), some researchers attempted to assess TR through the use of process measures, which inform on the quality of the therapeutic process by looking at some indicators, such as cohesion, alliance, and empathy. However, these instruments do not provide details on the therapists' actions taken to achieve the therapeutic objectives (Stiles & Wolfe, 2006). Thus, even though evaluating TR can be complex and expensive, it is a crucial need, as measuring TR could help therapists promote therapeutic success. Therefore, to fill this gap, the present systematic review aims to explore different methodologies (direct or indirect) of assessment, measurement, and analysis of TR in the field of psychotherapy research.

Methods

This systematic review was reported following the instructions provided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (Page *et al.*, 2021) (*Supplementary Figure 1*).

Information sources and research strategies

The research was conducted between February 2023 and April 2023 by two independent researchers (second and fourth authors), resorting to the following databases: APA PsycInfo, PubMed and Scopus. The selected keywords, searched in the full text, were the following: ("Therapist" OR "Psychotherapist" OR "Psychologist" OR "Clinician") AND ("Therapist Responsiveness") AND ("Evaluat*" OR "Assess*" OR "Measur*" OR "Analys*" OR "Analyze*") AND ("Tool" OR "Instrument" OR "Scale" OR "Questionnaire" OR "System" OR "Grid" OR "Inventory" OR "Index"). We selected the English language as the only filter, while we decided not to apply any time limiter regarding the publication of articles since, to our knowledge, the present systematic review is the first on the subject.

Selection and data collection procedures

The inclusion criteria of the present study were the following: i) measurement of TR through specific tools; and ii) research articles. Regarding the first criterion, studies were admitted both if they used instruments developed for a direct assessment of TR and if they employed instruments developed for the analysis of related constructs (*e.g.*, empathy, attunement), provided that it was explicitly stated that the measurement of the related construct constituted an indirect (substitute or inferential) measure of TR. Concerning the second criterion, it should be specified that studies were admitted only if they were empirical articles; therefore, we

excluded book chapters and theoretical articles. We decided not to define restrictive inclusion criteria regarding specific professionals, such as only psychotherapists, since the literature showed that TR is an interpersonal competence that can be developed in any clinical-psychological field.

Before proceeding with the screening phase, the duplicate articles from the three databases were removed. The selection then took place, at first, based on the title and abstract and, later, through the reading of the full text. These selection and data collection procedures were carried out by two researchers (the second and fourth authors) through consensus agreement, and a third researcher (the first author) resolved any disagreement as a judge. The information extracted for each study concerned the authors, year of publication, country of the research, definition of TR, constructs associated with TR (e.g., outcome or WA), research objectives, research design, sample size, instruments used for the assessment of TR, presence of additional instruments, results, and conclusions.

Quality assessment

The quality assessment of the eligible articles was carried out by two independent researchers (the second and fourth authors), through consensus agreement, using the following instruments: a

checklist for case reports (Moola *et al.*, 2020), a checklist for cohort studies (Moola *et al.*, 2020), the revised Joanna Briggs Institute (JBI) critical appraisal tool for assessment of risk of bias for randomized controlled trials (Barker *et al.*, 2023), and a checklist for qualitative research (Lockwood *et al.*, 2015).

Each one of these instruments provides the possibility to answer “yes”, “no”, “unclear”, or “not applicable”. Each item has received a score of “1” if it matched the specific criteria, or a score of “0” if the criteria were not incorporated, not clear, or not applicable. Finally, each article has been categorized into one of the three available classes, based on the rate of “yes” obtained: class A, of high quality, with at least 75% of affirmative answers; class B, of good quality, with a score of affirmative answers ranging from 51% to 74%; class C, of decent quality, with a score of affirmative answers less than 50%.

Results

As shown in Figure 1, the research produced a total of 329 articles, of which 26 were removed as duplicates. In the next step, 270 studies were excluded based on the title and abstract because they did not meet the inclusion criteria: in particular, 214 articles referred to other professional fields, such as physiotherapy and

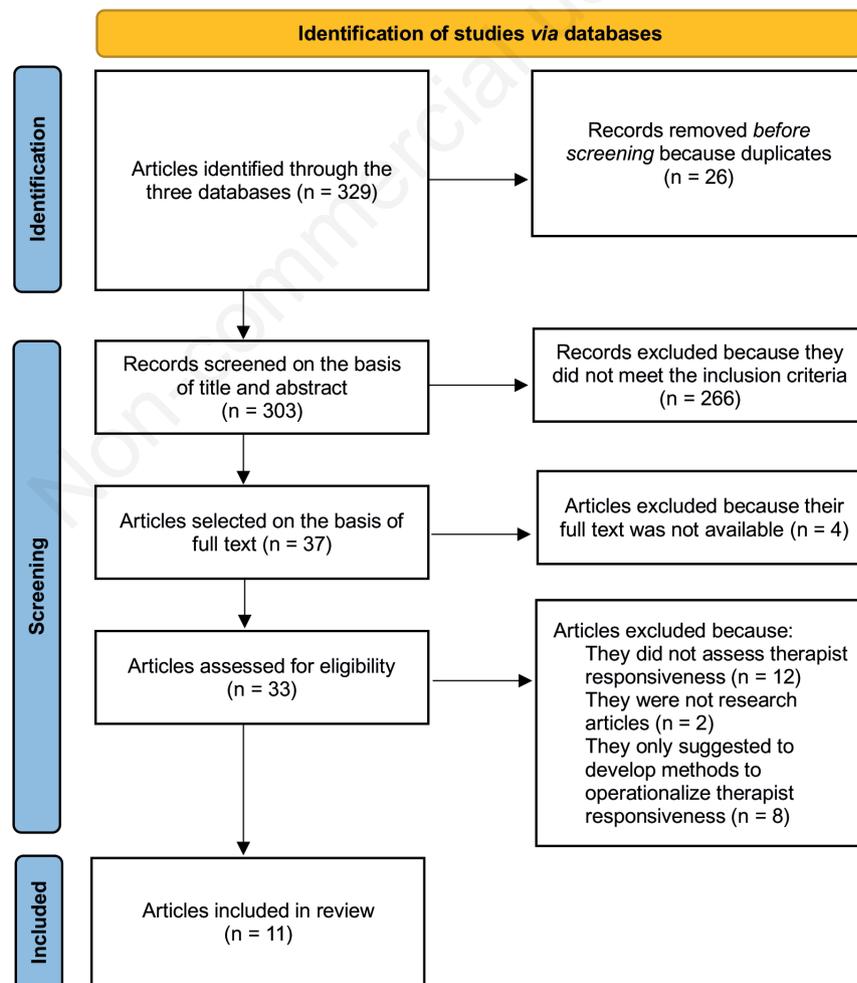


Figure 1. PRISMA 2020 flow diagram.

medicine; 31 studies focused on the theorization but not on the assessment of TR; 21 articles concerned occupational therapy, that is, rehabilitation for people with physical and/or cognitive disabilities through several recreational activities aimed at promoting autonomy in the tasks of daily life; 1 study was in German; 1 article was unavailable, despite having contacted the authors; finally, 2 studies were book chapters. As a result, the remaining 33 papers were considered for eligibility by reading the full text, either because they had already met the inclusion criteria or in an attempt to evaluate whether they met all the criteria since it was not clear from the title and abstract. Of these, 12 studies were excluded as they focused on the construct of TR as a means to open a discussion on those variables underlying the therapeutic change without mentioning the possibility of assessment of this construct; 2 articles were excluded, although they highlighted the importance of the assessment of TR in supervision, because they were not research articles and did not provide guidance on any instruments to be used; and finally, 8 articles were excluded as they only underlined the necessity for researchers to develop methods or theoretical models to operationalize TR. The remaining 11 articles were considered eligible as they met all the inclusion criteria, constituting the final pool of this systematic review. The full list of included articles and information extracted for each study is presented in *Supplementary Table 1*.

Summary of results

Concerning the definition of TR, 55% of included studies referred to the definition provided by Stiles *et al.* (1998) (Kivlighan *et al.*, 2017; Kramer *et al.*, 2014; Kramer *et al.*, 2016; Meystre *et al.*, 2014; Richards *et al.*, 2013; Spagnuolo Lobb *et al.*, 2022), while in the remaining studies, in 27%, the authors used the conceptualization by Hatcher (2015) (Snyder & Aafjes-van Doorn, 2016; Snyder & Silberschatz, 2016; Zalaznik *et al.*, 2021), and in 18%, the one by Elkin *et al.* (2014) (Culina *et al.*, 2022; Elkin *et al.*, 2014).

In relation to instruments employed to assess TR, in 55% of the included studies (Culina *et al.*, 2022; Elkin *et al.*, 2014; Richards *et al.*, 2013; Snyder & Aafjes-van Doorn, 2016; Snyder & Silberschatz, 2016; Zalaznik *et al.*, 2021) the authors adopted tools aimed at directly measure TR [*i.e.*, the therapist responsiveness scale (TRS), the patient's experience of attunement and responsiveness (PEAR) scale, and the adherence responsiveness measure], while in the remaining 45%, TR was assessed indirectly, through: i) the use of instruments to measure related constructs [*e.g.*, aesthetic relational knowledge scale (ARKS) or processing-content-relationship scale] (Kramer *et al.*, 2016; Spagnuolo Lobb *et al.*, 2022); ii) the integration of different constructs and methodologies, as in the case of discrepancy between working alliance inventory short form and real relationship inventory, or in the case of co-presence of the comprehensive psychotherapeutic interventions rating scale (C-PIRS) and the assimilation of problematic experiences scale (APES) (Kivlighan *et al.*, 2017; Meystre *et al.*, 2014); iii) the adoption of an atheoretical approach (*i.e.*, the motive-oriented therapeutic relationship) (Kramer *et al.*, 2014). Among the included studies, the only instruments used in more than one article were the PEAR scale (27%) (Snyder & Aafjes-van Doorn, 2016; Snyder & Silberschatz, 2016; Zalaznik *et al.*, 2021) and the TRS (18%) (Culina *et al.*, 2022; Elkin *et al.*, 2014).

Regarding the sources of information, 55% of the included studies were based on self-reported instruments. In 9% of these, instruments were administered only to therapists (Spagnuolo Lobb

et al., 2022); in 9%, only to patients (Meystre *et al.*, 2014); and in 37%, to both therapists and patients (Kivlighan *et al.*, 2017; Snyder & Aafjes-van Doorn, 2016; Snyder & Silberschatz, 2016; Zalaznik *et al.*, 2021). In 45% of included studies, measurement was based on observer-rated tools; in particular, 36% of the included studies used videotapes (Culina *et al.*, 2022; Elkin *et al.*, 2014; Kramer *et al.*, 2016; Richards *et al.*, 2013), and 9% analyzed transcripts of sessions (Kramer *et al.*, 2014).

Referring to the associated constructs, in 45% of the articles, TR has been related to the outcome (*i.e.*, symptom relief of panic disorders, depression, and personality disorders), finding a positive association in all studies (Kramer *et al.*, 2014; Kramer *et al.*, 2016; Richards *et al.*, 2013; Snyder & Silberschatz, 2016; Zalaznik *et al.*, 2021). In the remaining studies, TR has been related to therapist expertise in training (Spagnuolo Lobb *et al.*, 2022), quality session evaluation (Kivlighan *et al.*, 2017), assimilation of problematic experiences (Meystre *et al.*, 2014), and patient feedback (Snyder & Aafjes-van Doorn, 2016). Lastly, TR was also found to be positively associated with WA (Culina *et al.*, 2022) and early patient engagement (Elkin *et al.*, 2014).

Finally, considering the quality assessment, most of the 11 studies included in this systematic review are in class B (73%), 18% of articles are in class A, and only 9% of the studies are in class C (Table 1).

Discussion

This systematic review has synthesized data from empirical studies to explore the different methodologies of assessment of TR. Before the discussion of the results regarding our research question, we will comment on the results regarding the features of the included studies.

First of all, regarding the definition of TR, it is possible to note that the majority of included studies adopted the conceptualization by Stiles *et al.* (1998), in which TR represents a therapist's interpersonal competence consisting in a set of therapist's behaviors that are influenced by the emerging therapeutic context. The remaining studies reported the definition elaborated by Hatcher (2015), who considers TR as the therapist's ability to gain positive effects for the patient by adapting interventions to the current state of the patient and therapeutic process, or the formulation by Elkin *et al.* (2014), according to which TR indicates the extent to which the therapist appears caring, respectful, and interested in the patient's communications, both in terms of

Table 1. Results of the quality assessment of the included studies.

Study	Quality assesment, %	Level
Richards <i>et al.</i> (2013)	69	B
Elkin <i>et al.</i> (2014)	62	B
Kramer <i>et al.</i> (2014)	53	B
Meystre <i>et al.</i> (2014)	90	A
Snyder & Aafjes-van Doorn (2016)	88	A
Snyder & Silberschatz (2016)	53	B
Kramer <i>et al.</i> (2016)	60	B
Kivlighan <i>et al.</i> (2017)	62	B
Zalaznik <i>et al.</i> (2021)	60	B
Culina <i>et al.</i> (2022)	62	B
Spagnuolo Lobb <i>et al.</i> (2022)	50	C

content and feelings. This result confirms the lack of a unique definition of TR, hindering its operationalization and, consequently, its measurement.

It is also possible to observe that all the included studies concerned the psychotherapeutic field, despite our attempt to collect data in relation to the wider psychological-clinical field. This result suggests that future research should broaden the context of the investigation and application of TR.

Regarding the sources of information, most of the included studies adopted observer-rated tools, especially those based on videotapes. Considering that TR includes non-verbal communication, this type of analysis could be more informative than those based on transcripts, as it allows for the consideration of several levels of TR. The remaining studies employed self-reported instruments, in particular those administered to both therapists and patients. In this last case, a great discrepancy emerged between therapists' and patients' ratings, underscoring the subjective characteristic of TR. Indeed, as stated above, therapists could judge their intervention as responsive due to its consistency with guidelines provided by their own theoretical model, but patients could consider it not responsive based on their experience, for example, if it was distressing (Kivlighan *et al.*, 2017).

With reference to the relationship between TR and associated constructs, most of the included studies explored the impact of TR on treatment outcomes, finding a positive association between these variables. Other studies showed that TR may positively influence outcomes indirectly, particularly through a positive impact on patient engagement and WA. Thus, most of the studies included in the current systematic review emphasize the relevance of TR in favoring treatment effectiveness.

Moving to the main research question that motivated this systematic review, namely "which type of (direct or indirect) instruments are used to assess the therapist responsiveness?", a great heterogeneity emerged in the choice and the use of the tools. In particular, we could detect two trends in terms of methods of measurement. The first one, found in the majority of the included studies, assessed TR using direct and explicit tools. The TRS (Elkin & Smith, 2007) is an observer-rated tool developed for the assessment of TR and subdivided into three parts, and each item is rated on a Likert scale ranging from 0 to 4. Specifically, part I assesses 5-minute intervals of sessions and measures the presence of appropriate responsiveness as well as behaviors indicating a lack of appropriate responsiveness; part II considers the whole session and focuses on the overall therapeutic atmosphere, to which both therapist and patient contribute; and part III aims to obtain the rater's global impression of the therapist and patient. Particularly relevant is the global responsiveness item included in the third part, which consists of a global rating of TR. Elkin *et al.* (2014) showed the reliabilities of all responsiveness factors: attentiveness (Cronbach's $\alpha=.751$), early emphatic responding (Cronbach's $\alpha=.726$), negative therapist behavior (Cronbach's $\alpha=.697$), positive therapeutic atmosphere (Cronbach's $\alpha=.897$).

In this regard, it is important to clarify that, in the studies synthesized in the current systematic review, this tool has been adopted in different ways: Elkin *et al.* (2014) administered all three subscales, while Culina *et al.* (2022) applied only the third one, *i.e.*, the global responsiveness item, leaving out the other two components of the scale that could be relevant. However, it should be pointed out that in the study by Elkin *et al.* (2014), the third subscale appeared more significant in the assessment of TR. Anyway, there is no doubt that in the future, this tool may be a reference point for the direct assessment of this construct.

The PEAR scale (Silberschatz, 2009) is a self-report measure

developed to assess the patient's experience of the therapist's degree of attunement and responsiveness during the session. This scale has a patient version (PEAR-p) and a therapist version (PEAR-t), each containing 30 items, which are similar in content but rephrased to be appropriate for the patient or for the therapist (for example, PEAR-p item 8: "what my therapist did and said was helpful today", PEAR-t item 8: "what I did and said was helpful to my patient today"). Each item is rated on a Likert scale ranging from 0 (not at all) to 3 (very much), and then these responses are summed to achieve a total attunement and responsiveness score. PEAR-p scale demonstrated good reliability ($\omega=.94$) as did the PEAR-t ($\omega=.96$) (Snyder & Silberschatz, 2016). The present review also detected significant differences among the three studies in the use of the PEAR scale: in the first study, the authors used it to understand the impact of patients' experiences of attunement and responsiveness during psychotherapy sessions (Snyder & Silberschatz, 2016); in the second study, this tool was administered in an online context, requiring several changes compared to original items to adapt it to this different setting (Zalaznik *et al.*, 2021); and finally, in the third study, the authors employed this instrument in a case study as a feedback tool (Snyder & Aafjes-van Doorn, 2016).

The responsiveness adherence measure (Richards *et al.*, 2013) is an observer-rated scale composed of five items that assess aspects of TR using a nominal, dichotomous scale ("present" or "absent"). The five items and their reliability are the following: i) empathy and acknowledgment of emotions ($k=.484$); ii) provision of guidance and information ($k=.647$); iii) validation of participant's successes ($k=.478$); iv) promotion of self-care and social support ($k=.478$); and v) building alliance ($k=.657$). If all five aspects are present, then the score would be 5. This instrument was used in an online context to analyze the impact of the therapist in an e-cognitive behaviour therapy (e-CBT) (Richards *et al.*, 2013). In this regard, it is necessary to note that in such cases, the interactions between therapists and clients were asynchronous. However, as stated above, TR has a dynamic and moment-by-moment nature that is difficult to capture in an asynchronous relationship.

The second trend, found in the remaining 45% of studies, indirectly assessed TR. The ARKS is a self-report measure composed of 58 items evaluating empathy, resonance, and bodily awareness that can be administered only to therapists (Spagnuolo Lobb *et al.*, 2022). Participants are asked to respond on a 7-point Likert scale ranging from 1 (extremely disagree) to 7 (extremely agree). The reliability of the whole scale was good (Cronbach's $\alpha=.87$), and for the single dimensions (reliability of empathy=.672; resonance=.730; body awareness=.921) (Spagnuolo Lobb *et al.*, 2022).

In this regard, it is opportune to specify that the ARKS allows for the assessment of empathy, resonance, and bodily awareness only from a therapist's point of view, while TR is an interpersonal competence that needs to be measured within a specific clinical relationship due to its relational nature. This is the case of the processing-content-relationship scale (Sachse *et al.*, 2011), which takes into account the therapeutic interaction: indeed, it is an observer-rated instrument that evaluates the quality of the therapeutic interaction through 54 items, rated on a Likert-scale ranging from 0 (not at all) to 6 (to a great extend). Overall ratings are made for both the patient's and therapist's processes that occur for 10 minutes, using video/audio recordings of a mid-session segment (between minutes 10 and 20). Higher scores reflect better interaction quality. This tool showed good reliability (κ varied between .72 and .85). From the patient's perspective, three sub-scales are defined (content, process, and relationship), while from the thera-

pist's perspective, six sub-scales are defined (relationship, understanding, process directivity, treatment of the patient's avoidance, treatment of interactional maneuvers, treatment of schemes).

As indicated by Kivlighan *et al.* (2017), TR can also be assessed as a marker of discrepancy between the WA and the real relationship (RR), since a responsive therapist would emphasize a different aspect of the therapeutic relationship, *i.e.*, work-focused (WA) or authenticity-focused (RR), with each client depending on the client's needs. In this case, the assessment of TR occurs through the use of two instruments. The first is the working alliance inventory - short form (Hatcher & Gillaspay, 2006), which is a 12-item tool administered to both patient and therapist that allows for the evaluation of tasks, goals, and bonds. Items are scored on a 5-point scale, from 1 (seldom) to 5 (always). This tool showed good internal consistency both for the client form ($\alpha=.95$) and for the therapist form ($\alpha=.94$) (Kivlighan *et al.*, 2017). The second tool is the real relationship inventory (Gelso *et al.*, 2005), which is a 12-item instrument administered to both patient and therapist that assesses perceptions of genuineness and realism in the therapeutic relationship. Items are rated using a 5-point scale from 1 (strongly disagree) to 5 (strongly agree). This tool showed good internal consistency both for the client form ($\alpha=.90$) and for the therapist form ($\alpha=.90$) (Kivlighan *et al.*, 2017). Kivlighan *et al.* (2017) related these scores to the evaluation of session quality, showing that patients evaluated a session as higher in quality when they perceived that one or the other aspect (work-focused or authenticity-focused) of the relationship predominated (*i.e.*, in case of greater discrepancy), in contrast to therapists' perspectives, in which higher quality was associated with the perception of balance between the two aspects (*i.e.*, in case of lower discrepancy).

Another indirect way to evaluate TR requires the integration of other two instruments: the C-PIRS (Trijsburg *et al.*, 2002), which provides an exhaustive classification of therapeutic interventions, and the APES (Stiles, 2002), which indicates the level of assimilation (*i.e.*, integration of problematic experiences during psychotherapy) before and after the therapist's intervention, exploring what the latter has produced. In this way, it is possible to compare the APES levels for the patient's statements before and after each therapist's intervention coded according to C-PIRS in order to see which intervention is associated with APES progress/regression at each level. In the present review, one study integrated C-PIRS and APES, analyzing time to time the impact of therapists' interventions on the relationship with patients (Meystre *et al.*, 2014). Finally, TR was evaluated by a therapeutic approach, the motive-oriented therapeutic relationship as the degree of complementarity between the therapeutic approach and patients' plan analysis (Kramer *et al.*, 2014). The latter allows the therapist to explore the patient's needs to be able to provide individualized therapy that will not reinforce problematic elements. In this case, through transcripts and videotapes, observer-raters extrapolated portions of therapeutic interactions considered significant as inherent to the patient's plan analysis, rating them on a scale ranging from 1 to 7. This is particularly interesting because it differs from previous methods and instruments; therefore, future research should deepen this modality to clarify if it is useful in the assessment of TR.

Ultimately, concerning quality assessment, most of the included studies showed limited rigor about the scientific reporting of methodology, since a medium level of quality was found and only two studies presented a high-quality evaluation. However, Siddaway *et al.* (2019) affirm that the results of quality assessment should not be used to determine the exclusion of articles from review nor to evaluate the quality of the review, but exclu-

sively to take into account the potential risk of bias when drawing conclusions.

Conclusions

The present systematic review aimed at exploring the different methodologies of assessment of TR. The results of this study highlighted difficulties and divergences in the operationalization of the construct, as shown by the great heterogeneity found in the choice and use of the instruments. In this regard, it is possible to observe the scarcity of the use of direct tools, which in turn have also been administered in different ways, making it even harder to obtain comparable results, thus hindering the possibility of understanding the impact of TR on the therapeutic relationship and treatment effectiveness. Furthermore, some authors adopted observer-rated tools, while others preferred self-report measures. The latter undoubtedly provides relevant information, especially when assessing both therapists' and clients' perspectives. However, they are also strongly affected by participants' subjectivity. In fact, as stated above, therapists could define responsiveness as an intervention, but this intervention could be painful for patients, leading them to consider the therapist as not responsive. This aspect is underlined by the great discrepancy found between the therapists' and clients' ratings. In order to overcome the limitations of these instruments, it could be useful to triangulate the sources of information. In particular, analyses based on transcripts of the session (see Gelo *et al.*, 2021) could offer advantages since they provide information related to the interaction between the therapist and the client. However, TR includes non-verbal communication; therefore, analyses based on videotapes seem to be more suitable since they seem able to capture moment by moment the evolution of the therapeutic relationship.

In an attempt to identify the most appropriate methodology of assessment, it is important to reiterate that TR is declined within the clinical relationship and thus has a relational nature; therefore, we believe that it is contradictory to measure it without taking into account the interactions between therapist and client. Therefore, it would be advisable to avoid measuring such a construct in the asynchronous relationship or examining only the therapist's point of view, as if TR was a therapist's personal characteristic; on the contrary, TR is an interpersonal competence that is articulated within the specific therapeutic relationship.

Another interesting point is the relationship between TR and related constructs, such as therapist expertise (Spagnuolo Lobb *et al.*, 2022) and assimilation of problematic experiences (Meystre *et al.*, 2014), which are often wrongly overlapped without taking into account differences. It is necessary that future research clarify the definition and operationalization of TR in order to distinguish it from other variables and give mental health professionals indications for clinical practice. To this end, a pluralistic approach involving quantitative, qualitative, and mixed methods could be useful (Gelo *et al.*, 2015).

The present systematic review may have some implications for research, clinical practice, and the training of therapists. First of all, it guides researchers in choosing and developing direct tools to evaluate TR, suggesting paying particular attention to the relational and interactive nature of this construct. In addition, it provides an overview of tools that allow for the measurement of TR in a reliable way. Furthermore, it invites researchers to adopt a multi-informant and multi-method approach in the assessment of TR, triangulating the sources of information and the methodologies, such as therapists' and

clients' self-report, and observer-rated instruments based on videotapes of the sessions. This may enable more objective ratings and capture non-verbal aspects of communication. From a clinical practice perspective, this review suggests that therapists should monitor and be careful of their levels of responsiveness, looking also at the clients' perspective, to reduce discrepancies and adapt their interventions to the patient's needs. Finally, in terms of training, this study recommends including the promotion of TR in training procedures since it is an essential interpersonal competence able to positively affect the WA and treatment effectiveness (see Messina *et al.*, 2018). However, this systematic review presents some limitations. First of all, the small number of studies examined reduces the possibility of generalizing the results obtained. In addition, the identification and selection of studies were carried out through manual searches without the use of electronic tools, so it is possible that some articles were accidentally excluded. Lastly, despite both search and quality assessment being carried out by two independent researchers, they were not blinded, and this may have affected the evaluation of the included papers. However, such a revision is also characterized by several strengths. Firstly, it filled the gap in psychotherapy research, namely the lack of a synthesis of the literature on the methodologies of assessment of TR. Secondly, three databases were used to obtain a broader and more diversified pool of articles.

References

- Barker, T. H., Stone, J. C., Sears, K., Klugar, M., Tufanaru, C., Leonardi-Bee, J., Aromataris, E., & Munn, Z. (2023). The revised JBI critical appraisal tool for the assessment of risk of bias for randomized controlled trials. *JBI Evidence Synthesis*, 21(3), 494-506. doi: 10.11124/JBIES-22-00430.
- Cahill, J., Barkham, M., Hardy, G., Gilbody, S., Richards, D., Bower, P., & Connell, J. (2009). A review and critical appraisal of measures of therapist-patient interactions in mental health settings. *Health Technology Assessment*, 12(24), iii. doi: 10.1108/cgij.2009.24814aaf.003.
- Culina, I., Fiscalini, E., Martin-Soelch, C., & Kramer, U. (2022). The first session matters: therapist responsiveness and the therapeutic alliance in the treatment of borderline personality disorder. *Clinical Psychology & Psychotherapy*, 30(1), 131-140. doi: 10.1002/cpp.2783.
- Elkin, I., & Smith, Y. (2007). *Therapist responsiveness study rating scale*. School of Social Service Administration, University of Chicago, Chicago, Illinois.
- Elkin, I., Falconnier, L., Smith, Y., Canada, K. E., Henderson, E., Brown, E. R., & McKay, B. M. (2014). Therapist responsiveness and patient engagement in therapy. *Psychotherapy Research*, 24(1), 52-66. doi: 10.1080/10503307.2013.820855.
- Erskine, R. G. (1998). Attunement and involvement: therapeutic responses to relational needs. *International Journal of Psychotherapy*, 3(3), 235-244.
- Esposito, G., Di Maro, A., & Passeggia, R. (2024). The relationship between treatment integrity and outcome in group psychotherapy: A systematic review. *Clinical Psychology & Psychotherapy*, 31(1), 1-17. doi: 10.1002/cpp.2952
- Esposito, G., Passeggia, R., Cutolo, A. S., Karterud, S., & Freda, M. F. (2020). Treatment integrity and members' change in group counseling: A pilot study on counselor's mentalizing interventions. *Professional Psychology: Research and Practice*, 51(6), 588-597. doi: 10.1037/pro0000304
- Eubanks, C. F., Lubitz, J., Muran, J. C., & Safran, J. D. (2019). Rupture resolution rating system (3RS): development and validation. *Psychotherapy Research*, 29(3), 306-319. doi: 10.1080/10503307.2018.1552034.
- Eubanks, C. F., Sergi, J., & Muran, J. C. (2021). Responsiveness to ruptures and repairs in psychotherapy. In Watson, J. C., & Wiseman, H. (Eds.), *The responsive psychotherapist: attuning to clients in the moment* (pp.83-103). American Psychological Association.
- Gelo, O. C. G., Pritz, A., & Rieken, B. (2015). Preface. In O. C. G. Gelo, A. Pritz, & B. Rieken (Eds.), *Psychotherapy research: Foundations, process, and outcome*. Springer.
- Gelo, O. C. G., Salcuni, S., & Colli, A. (2012). Text analysis within quantitative and qualitative psychotherapy process research: Introduction to special issue. *Research in Psychotherapy: Psychopathology, Process and Outcome*, 15(2), 45-53. https://doi.org/10.7411/RP.2012.005
- Gelso, C. J., Kelley, F. A., Fuertes, J. N., Marmarosh, C., Holmes, S. E., Costa, C., & Hancock, G. R. (2005). Measuring the real relationship in psychotherapy: initial validation of the therapist form. *Journal of counseling psychology*, 52(4), 640-649. doi: 10.1037/0022-0167.52.4.640.
- Hatcher, R. L. (2015). Interpersonal competencies: responsiveness, technique, and training in psychotherapy. *American Psychologist*, 70(8), 747-757. doi: 10.1037/a0039803.
- Hatcher, R. L. (2021). Responsiveness, the relationship, and the working alliance in psychotherapy. In Watson, J. C., & Wiseman, H. (Eds.), *The responsive psychotherapist: attuning to clients in the moment* (pp. 37-58). American Psychological Association.
- Hatcher, R. L., & Gillaspay, J. A. (2006). Development and validation of a revised short version of the working alliance inventory. *Psychotherapy research*, 16(1), 12-25. doi: 10.1080/10503300500352500.
- Kendall, P. C., & Frank, H. E. (2018). Implementing evidence-based treatment protocols: flexibility within fidelity. *Clinical Psychology: Science and Practice*, 25(4), e12271. doi: 10.1111/cpsp.12271.
- Kivlighan, Jr, D. M., Kline, K., Gelso, C. J., & Hill, C. E. (2017). Congruence and discrepancy between working alliance and real relationship: Variance decomposition and response surface analyses. *Journal of Counseling Psychology*, 64(4), 394-409. doi: 10.1037/cou0000216.
- Kramer, U., Flückiger, C., Kolly, S., Caspar, F., Marquet, P., Despland, J. N., & De Roten, Y. (2014). Unpacking the effects of therapist responsiveness in borderline personality disorder: motive-oriented therapeutic relationship, patient in-session experience, and the therapeutic alliance. *Psychotherapy and Psychosomatics*, 83(6), 386-387. doi: 10.1159/000365400.
- Kramer, U., Pascual-Leone, A., Rohde, K. A., & Sachse, R. (2016). Emotional processing, interaction process, and outcome in clarification-oriented psychotherapy for personality disorders: a process-outcome analysis. *Journal of Personality Disorders*, 30(3), 373-394. doi: 10.1521/pedi_2015_29_204.
- Lockwood, C., Munn, Z., & Porritt, K. (2015). Qualitative research synthesis: methodological guidance for systematic reviewers utilizing meta-aggregation. *International Journal of Evidence-Based Healthcare*, 13(3), 179-187. doi: 10.1097/XEB.0000000000000062.
- Marques, L., Valentine, S. E., Kaysen, D., Mackintosh, M. A., Dixon De Silva, L. E., Ahles, E. M., Youn, S. J., Shtasel, D. L., Simon, N. M., & Wiltsey-Stirman, S. (2019). Provider fidelity and modifications to cognitive processing therapy in a

- diverse community health clinic: Associations with clinical change. *Journal of Consulting and Clinical Psychology*, 87(4), 357-369. doi: 10.1037/ccp0000384.
- McAleavey, A. A., & Castonguay, L. G. (2014). The process of change in psychotherapy: common and unique factors. In O. C. G. Gelo, A. Pritz, & B. Rieken (Eds.), *Psychotherapy research: foundations, process, and outcome* (pp. 293-310). Springer Vienna.
- Messina, I., Gelo, O. C. G., Sambin, M., Bianco, F., Mosconi, A., Fenelli, A., Curto, M., Gullo, S., & Orlinsky, D. (2018). Trainees' self-evaluation of their development as psychotherapists: An Italian contribution to an international collaborative study on psychotherapy training. *Clinical Psychology & Psychotherapy*, 25(2), 338-347. <https://doi.org/10.1002/cpp.2165>
- Meystre, C., Kramer, U., De Roten, Y., Despland, J. N., & Stiles, W. B. (2014). How psychotherapeutic exchanges become responsive: a theory-building case study in the framework of the assimilation model. *Counselling and Psychotherapy Research*, 14(1), 29-41. doi: 10.1080/14733145.2013.782056.
- Moola S., Munn Z., Tufanaru C., Aromataris E., Sears K., Sfectu R., Currie M., Qureshi R., Mattis P., Lisy K., & Mu P-F. (2020). Chapter 7: Systematic reviews of etiology and risk. In: Aromataris E., & Munn Z (Eds.), *JBIManual for Evidence Synthesis*. JBI.
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., & McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *British Medical Journal*, 372, n160. doi: 10.1136/bmj.n160.
- Reis, H. T. (2014). Responsiveness: affective interdependence in close relationships. In M. Mikulincer & P. R. Shaver (Eds.), *Mechanisms of social connection: from brain to group* (pp. 255-271). American Psychological Association. doi: 10.1037/14250-015.
- Richards, D., Timulak, L., & Hevey, D. (2013). A comparison of two online cognitive-behavioural interventions for symptoms of depression in a student population: the role of therapist responsiveness. *Counselling and Psychotherapy Research*, 13(3), 184-193. doi: 10.1080/14733145.2012.733715.
- Sachse, R., Schülken, T., Leisch, M., & Sachse, M. (2011). Effektivität Klärungsorientierter Psychotherapie: erste Ergebnisse. In R. Sachse, J. Fasbender, J. Breil, & M. Sachse (Eds.), *Perspektiven Klärungsorientierter Psychotherapie II* (pp. 239-276). Lengerich: Pabst Science Publishers. [Material in German].
- Siddaway, A. P., Wood, A. M., & Hedges, L. V. (2019). How to do a systematic review: a best practice guide for conducting and reporting narrative reviews, meta-analyses, and meta-syntheses. *Annual review of psychology*, 70, 747-770. doi: 10.1146/annurev-psych-010418-102803.
- Silberschatz, G. (2009). *The patient's experience of attunement and responsiveness scale* (Unpublished manuscript).
- Snyder, J., & Aafjes-van Doorn, K. (2016). Utilizing measure-based feedback in control-mastery theory: a clinical error. *Psychotherapy*, 53(3), 291-296. doi: 10.1037/pst0000075.
- Snyder, J., & Silberschatz, G. (2016). The patient's experience of attunement and responsiveness scale. *Psychotherapy Research*, 27(5), 608-619. doi: 10.1080/10503307.2016.1147658.
- Spagnuolo Lobb, M., Sciacca, F., Iacono Isidoro, S., & Di Nuovo, S. (2022). The therapist's intuition and responsiveness: What makes the difference between expert and in training Gestalt psychotherapists. *European Journal of Investigation in Health, Psychology and Education*, 12(12), 1842-1851. doi: 10.3390/ejihpe12120129.
- Stiles, W. B. (2002). Assimilation of problematic experiences. In J. C. Norcross (Eds.), *Psychotherapy relationships that work: therapist contributions and responsiveness to patients* (pp. 357-365). Oxford University Press.
- Stiles, W. B. (2021). Responsiveness in psychotherapy research: problems and ways forward. In Watson, J. C., & Wiseman, H. (Eds.), *The responsive psychotherapist: attuning to clients in the moment* (pp. 15-35). American Psychological Association. doi: 10.1037/0000240-002.
- Stiles, W. B., & Horvath, A. O. (2017). Appropriate responsiveness as a contribution to therapist effects. In L. G. Castonguay & C.E. Hill (Eds.), *How and why are some therapists better than others? Understanding therapist effects* (pp. 71-84). American Psychological Association. doi: 10.1037/0000034-005.
- Stiles, W. B., & Wolfe, B. E. (2006). Relationship factors in treating anxiety disorders. In L. G. Castonguay & L. E. Beutler (Eds.), *Principles of therapeutic change that work* (pp. 155-165). Oxford University Press.
- Stiles, W. B., Honos-Webb, L., & Surko, M. (1998). Responsiveness in psychotherapy. *Clinical Psychology: Science and Practice*, 5(4), 439-458. doi: 10.1111/j.1468-2850.1998.tb00166.x.
- Trijnsburg, R. W., Frederiks, G. C. F. J., Gorlee, M., Klouwer, E., Hollander, A. D., & Duivendoorn, H. J. (2002). Development of the comprehensive psychotherapeutic interventions rating scale (CPIRS). *Psychotherapy Research*, 12(3), 287-317. doi: 10.1093/ptr/12.3287.
- Watson, J. C., & Wiseman, H. (Eds.). (2021). *The responsive psychotherapist: attuning to clients in the moment*. American Psychological Association. doi: 10.1037/0000240-000.
- Zalaznik, D., Strauss, A. Y., Halaj, A., Barzilay, S., Fradkin, I., Katz, B. A., & Huppert, J. D. (2021). Patient alliance with the program predicts treatment outcomes whereas alliance with the therapist predicts adherence in internet-based therapy for panic disorder. *Psychotherapy Research*, 31(8), 1022-1035. doi: 10.1080/10503307.2021.1882712.
- Zuroff, D. C., Kelly, A. C., Leybman, M. J., Blatt, S. J., & Wampold, B. E. (2010). Between-therapist and within-therapist differences in the quality of the therapeutic relationship: effects on maladjustment and self-critical perfectionism. *Journal of Clinical Psychology*, 66(7), 681-697. doi: 10.1002/jclp.20683.

Online supplementary material:

Supplementary Table 1. Characteristics of the included studies.

Supplementary Figure 1. PRISMA 2020 checklist.