Toward Personalized Psychotherapy: The Importance of the Trait-Like/State-Like Distinction for Understanding Therapeutic Change

Sigal Zilcha-Mano


Toward Personalized Psychotherapy: The Importance of the Trait-Like/State-Like Distinction for Understanding Therapeutic Change

Sigal Zilcha-Mano
University of Haifa

For the past hundred years, mechanisms of change have been the black box of psychotherapy. Thousands of studies failed to produce consistent findings, even concerning factors considered crucial for treatment success by theoretical models and decades of clinical experience. This article introduces the distinction between trait-like (TL) and state-like (SL) components of any mechanism of change (the TLSL distinction) as a potential key to the black box of psychotherapy. TL refers to individual differences between patients; SL refers to changes occurring within the patient over the course of treatment. The TLSL distinction explains why past research, which conflated the two, has produced conflicting results, and predicts the conditions under which consistent results can be obtained. Data collected so far show support for the importance of the TLSL distinction and point the way toward personalized treatment. The TL components create the individual’s signature pathology and strengths map, and determine the SL changes that represent the patient-specific mechanisms most critical for optimizing treatment efficacy for each individual. The TLSL distinction has the potential to explain not only how psychotherapy works, but also how changes of any type occur in the wake of intervention, life events, and other factors.

Public Significance Statement
The article introduces the distinction between trait-like (TL) and state-like (SL) components of each construct in psychotherapy (the TLSL distinction) as a potential key to the black box of psychotherapy. TL refers to individual differences between patients, SL to changes occurring within patients over the course of treatment. The TLSL distinction may pave the road to optimizing treatment efficacy. The patients’ main TL strengths and weaknesses can serve as a map identifying the SL changes each one needs, which then can serve as the basis for designing accurate individual treatment plans for each patient. Once the black box of psychotherapy lies open, the promise of evidence-based, personalized treatment can be fulfilled.

Keywords: mechanisms of change, therapeutic alliance, psychotherapy process research

Supplemental materials: http://dx.doi.org/10.1037/amp0000629.supp

After a century of theoretical and clinical writing in psychotherapy, and more than 80 years of research, many important mechanisms of change have been identified, but few consistent findings have accumulated about what makes treatment work (Cuijpers, Reijnders, & Huibers, 2019). One of the few consistent finding reported on mechanisms of change is that a stronger therapeutic alliance is associated with better treatment outcomes (Barber, 2009; Crits-Christoph, Gibbons, & Mukherjee, 2013; Flückiger, Del Re, Wampold, & Horvath, 2018). This finding is so robust, that according to a recently published meta-analysis, 1,000 studies claiming null effect would be needed to challenge it (Flückiger et al., 2018). We also know that almost every treatment works for about 50% of patients, and that different treatments, based on distinct mechanisms, are about equally effective (Cuijpers et al., 2014). Without understanding how treatment works, we can do little to improve it (Kazdin, 2014). A key to the black box of psychotherapy may lie in other fields of science (e.g., statistics, medicine), where an important distinction was made between what we refer to here as trait-like (TL) and state-like (SL) components of mechanisms of change, or the TLSL distinction.

The writing of this article was supported by two grants from the Israeli Science Foundation (Grant 186/15; 395/19).

Correspondence concerning this article should be addressed to Sigal Zilcha-Mano, Department of Psychology, University of Haifa, Mount Carmel, Haifa 31905, Israel. E-mail: sigalzil@gmail.com
The 100-Year-Old Question

That therapy works for some people some of the time became obvious from the start. The question that remains the black box of psychotherapy is how treatment works. This question has been addressed best by research on mechanism of change, or more broadly, on process variables. Process variables include events during psychotherapy sessions, or constructs thought to change during or between sessions as a consequence of therapeutic interactions, which subsequently lead to changes in presenting problems, symptoms, and functioning (Crits-Christoph et al., 2013). This definition refers to components that have the capacity to change over time, and whose change can lead to subsequent changes in outcome. Effective psychotherapy, or any effective intervention for that matter, aimed at reducing any type of emotional suffering and improving mental health, requires knowledge about how treatment works. Elucidating the mechanisms that bring about therapeutic change can reveal whether given treatments work through similar or distinct mechanisms, and whether different individuals benefit from the same or different mechanisms (Kazdin, 2007). Knowing how treatment works for a given individual will enable us to understand which active components have a beneficial effect for that individual and which do not, opening the door to cost-effective personalized therapy. Identifying the most curative mechanisms for a given individual, on the one hand, and the concrete intervention (or group of techniques) that target each mechanism or active ingredient, on the other hand, can help create a better match between the two: each individual can receive the tailored interventions or techniques that are the most powerful facilitators of change for that individual. It will make it possible to use this knowledge to devise treatments that include exactly the ingredients that are useful to each patient, eliminating those that are harmful, and reducing those that are indifferent and take up valuable treatment time without benefiting the patient.

The content of the black box of psychotherapy has preoccupied humanity for more than 100 years (e.g., Freud, 1900). A great deal has been written about it, from poetry (Sexton, 1959) to manuals explaining how therapy works (Barlow et al., 2017; Luborsky & Crits-Christoph, 1998), to empirical research based on randomized controlled trials (RCTs). All sought to answer this question using a range of investigative tools. At first, these were descriptions of individual cases (e.g., the case of Dibs; Axline, 1964; “The Two Analyses of Mr. Z.”; Kohut, 1979). The pioneers of psychotherapy, starting with Freud, described what caused a given treatment to work, and based on it, constructed theories about curative factors in therapy (e.g., Kohut, 1984). Patients wrote books (e.g., Danquah, 1999; Petersen, 2017), poems (e.g., “The Hanging Man” poem, by Plath, 1981), and made drawings (e.g., Psychoanalyst by Marie-Louise Von Motesiczky, 1962) to express what they thought caused their therapy to work. When systematic research to identify the mechanisms that bring about therapeutic change began, each treatment theory and approach developed an operationalized conceptual model and examined whether these mechanisms were associated with therapeutic change, using a variety of research designs and specialized statistical methods, such as mediation models. Behavioral therapies sought the key in behavioral change (Cuijpers, Van Straten, & Warmerdam, 2007; Foa, Hembree, & Rothbaum, 2007), cognitive models in cognitive change (Beck & Beck, 1995), emotion-based models in emotion expression and regulation (Fosha, Siegel, & Solomon, 2009; Greenberg, 2015), and psychodynamic models in insight into repetitive maladaptive patterns (Crits-Christoph & Barber, 1991).

This massive enterprise aimed at prying open the black box was able to decipher precious little of its content. The vast investment of time, energy, and funds returned little consistent information about why therapy works. Although intensive efforts have been made to develop new psychotherapies and improve existing ones, the effect size of treatments, like those for major depressive disorder (MDD), has not increased in the past decades, and many patients do not respond to treatment or relapse soon after recovery (Cuijpers, 2017). Even in the case of MDD, the leading cause of disability worldwide (Friedrich, 2017), therapy was found to help only about half the people who seek treatment, regardless of which out of hundreds of available treatments the patient was referred to. Some 500 RCTs have revealed that a wide variety of treatments based on different mechanisms are equally effective, and as effective as drug therapy (Cuijpers, 2017). Even focusing on mechanisms that
were successfully guiding clinicians for years, such as gaining insight into one’s maladaptive repetitive patterns or schemas (Crits-Christoph et al., 2013), or achieving cognitive change (Kazdin, 2007; Lorenzo-Luaces, German, & DeRubeis, 2015), yielded mixed results. Summarizing these studies, Kazdin (2009) noted that despite the attention paid to mechanisms of change and the “rather vast literature, there is little empirical research to provide an evidence-based explanation of precisely why treatment works and how the changes come about” (p. 419). Some have even argued that “It is as if we have been in the pilot phase of research for five decades without being able to dig deeper” (Cuijpers et al., 2019, p. 224). As a result, many researchers have proposed to shelve all our theoretical models trying to explain what makes treatment effective, and concluded that any change in treatment is explained entirely by factors common to all treatments. Others have claimed that everything works for everyone, and some even argued that nothing works for anyone (Wampold & Imel, 2015).

**Picking the Lock**

Based on numerous methodological papers (e.g., Wang & Maxwell, 2015), we know that every construct measured over time contains a TL component (baseline individual characteristics) and an SL one (characteristics that develop over the course of treatment). Much has been written in the methodological literature about analytic strategies for disentangling between-individuals variance (TL effects) from within-individual variance (SL effects; Curran & Bauer, 2011; Molenaar, 2004; Wang & Maxwell, 2015). Such strategies include: (a) centering—subtracting from each individual’s measurements the mean of that individual’s measurements or the baseline levels, (b) detrending—removing the time trend in addition to centering, and (c) latent variables—removing a latent variable consisting of baseline variables contributing to the construct. Another approach uses the same construct to measure the TL and SL components, but with different instructions. Two common examples of this approach are the state and trait anxiety (Spielberger, 1983) and anger (Spielberger, Jacobs, Russell, & Crane, 1983) scales, where the same items are used to measure TL and SL with different instructions, asking patients how they feel in general (TL) or at the present moment (SL). Although this approach is quite common, it has elicited some criticism (Lance, Christie, & Williamson, 2019). The Working Alliance Inventory, administered during treatment to measure alliance with the therapist, or before the start of treatment to measure expectations of alliance (Barber et al., 2014), is another example of how this approach can be implemented. In addition to statistical strategies, the design of the study can be adjusted to capture baseline and in-treatment temporal patterns of the construct in focus (see a detailed guideline for study design and statistical strategies in the online supplemental materials).

Studies conducted in many fields of science show that trying to infer about the TL effect from the SL effect can lead to wrong conclusions (Curran & Bauer, 2011; Fisher, Medaglia, & Jeronimus, 2018; Wang & Maxwell, 2015), and that each effect may have a different, and even opposite association with the outcome variable. For example, empirical research in cardiology suggests that individuals are more likely to have a heart attack during vigorous exercise (SL effect). Yet, individuals who exercise more are at lower risk of heart attack (TL effect; Curfman, 1993). Clearly, we cannot infer about the TL effect from the SL effect, lest we reach the wrong conclusion that people should stop exercising. Preliminary evidence shows that in psychotherapy, TL and SL effects can have opposite associations with effectiveness of treatment (Zilcha-Mano, 2017), as demonstrated in Figure 1.

Figure 1 describes the potential effect of patient insight on treatment outcome. The association between insight and the efficacy of psychotherapy that focuses on improving insight may differ at the TL and SL levels. At the TL level, patients with poorer insight may benefit most from psychotherapy that focuses on improving insight. Thus, the association between insight and treatment efficacy at the TL level may be negative. At the SL level, however, the association may be positive, so that improving insight is expected to result in better efficacy. Again, the two effects may run in opposite directions, just like in the example of the association between exercise and heart attacks. Consider the following two patients: Ben, who started treatment with a high TL level of insight, and John, who started treatment with a low one. Both patients were asked to report interactions with significant others on the Self-Understanding of Interpersonal Patterns Scales–Interview (SUIP-I; Gibbons & Crits-Christoph, 2017), as administered pretreatment. John found it difficult to explain what exactly he expected from others while interacting with them, and was unable to cite any similarities between two interactions he reported, which were almost identical in the interpersonal conflicts apparent in them (as evaluated by two experienced independent evaluators). Ben, however, showed high levels of insight regarding his expectations from others in interpersonal relationships, as well as high levels of insight regarding adaptive and maladaptive tendencies he has in interpersonal relations. Although both patients may highly benefit from treatment that focuses on gaining insight into interpersonal relations, such as supportive-expressive treatment (Luborsky, 1984), it is reasonable to suggest that John shows greater potential to benefit from such treatment than Ben because at the TL level individuals with lower levels of insight may benefit more from treatment that focuses on improving insight than are individuals with higher levels of insight (a negative association). The SL effect, by contrast, is a
within-patient effect. According to the SL effect, gains in insight by both patients over the course of treatment are expected to be associated with subsequent gains in treatment efficacy (a positive association). Thus, although the TL effect is supposed to be negative, the SL effect is supposed to be positive, and the two effects may run in opposite directions. Scenarios of this type may be responsible for the mixed effect documented in the literature regarding the association between a variety of mechanisms of change, for example, insight, and outcome (Crits-Christoph et al., 2013).

Without the TLSL distinction, mechanisms of change appear to work consistently only when both SL and TL effects move in the same direction. This is what happens in the case of alliance, the only clearly consistent mechanism of change (Flückiger et al., 2018): those with better TL interpersonal abilities to create a stronger therapeutic alliance are more likely to benefit from treatment, and improvement in SL alliance during treatment is expected to improve subsequent treatment outcome (Zilcha-Mano, 2017).

The TLSL Distinction at Work

Consider expectancy. Extensive literature on concepts like self-fulfilling prophecies suggests that our expectations massively affect our functioning (Rosenthal, 2010). Yet, a recent meta-analysis based on 81 independent samples has found only a relatively small, often nonsignificant effect of expectancy on outcome ($r = .18$), characterized by great variation between studies (Constantino, Vislă, Coyne, & Boswell, 2018). Whereas in some of these studies expectancy had a significant effect on treatment outcome, in others it did not. The TLSL distinction may be instrumental in determining whether or not expectancy is a true mechanism of change. For example, in 128 patients receiving treatment for depression, where no TLSL distinction was made, the resulting effect was similar in size to that obtained in the aforementioned meta-analysis, and was not significant (Zilcha-Mano, Brown, Roose, Cappetta, & Rutherford, 2019). But the TLSL distinction, achieved by separating SL from TL using the standard statistical methods of disentangling the aggregated level (the TL component) and deviations from the aggregated level (the SL component), revealed that the effect of the TL component was not significantly different from 0. The SL effect on subsequent treatment outcome, however, was significant, and explained 4.6% of the variance in treatment outcome (Zilcha-Mano et al., 2019). This finding points to the great promise of the TLSL distinction. Once we know that the SL effect of expectancy is significant, we know that increasing expectancy increases the potential for subsequent symptomatic reduction. We can also identify the individuals for whom this is indeed the case, that is, for whom the SL effect on
treatment outcome is significant. Such individuals may be the ones whose TL tendencies show low expectancy, having therefore a clear deficit in this area. For these individuals, increasing expectancy is likely to result in subsequent improvements in treatment outcome. Thus, we can make use of the results of decades of research on, for example, self-fulfilling prophecy, to identify interventions that can be administered to patients before starting treatment (e.g., in the form of psychoeducation), to facilitate subsequent therapeutic change. Studies have shown that instructing students, or their teachers, that they have a good chance of success has indeed resulted in such success (Rosenthal, 2010). Similarly, to explore this mechanism in psychotherapy, in future studies patients or their therapists could receive, before treatment, information intended to increase their expectations regarding the success of the treatment, which ultimately may be associated with this very outcome (Constantino, Ametrano, & Greenberg, 2012).

Such manipulation of SL expectancy is critical for investigating whether outcome expectancy has the potential to be a mechanism of change. In one of the few studies that manipulated changes in SL expectancy before the beginning of antidepressant treatment, a one-sentence difference in the instructions given to patients to raise their SL expectancy resulted in profound differences between the groups in the amount of symptom reduction resulting from treatment. One group, the high expectancy condition, was told that they had a 100% chance of receiving antidepressant medication; the other, the low expectancy condition, was told they had a 50% chance of receiving medication and a 50% chance of receiving placebo. Individuals in the high expectancy condition showed greater increase in state-like expectancy and better outcome than did those in the low expectancy condition (Rutherford et al., 2017). The TL component in this example is the individuals’ general tendency toward high and positive expectations about the outcome of their treatment at baseline, before the start of treatment or the manipulation. This can be the product of many factors, including cultural assumptions about mental health treatment (Tseng, 2001), the individual’s general optimistic versus pessimistic tendencies (Seligman, 2002), internal representations of others as able and willing to help (Bowlby, 1988), and so forth. The SL component consists of changes in expectancy, following the SL expectancy manipulation. Rutherford et al. (2017) showed that changes in SL expectancy significantly mediated the effect of the expectancy manipulation on treatment outcome.

A follow-up study of Rutherford et al. (2017), investigating a potential neurobiological mechanism at the basis of the SL expectancy effect (Zilcha-Mano et al., 2019), focused on the hyperactivation of the amygdala in response to sad faces, which is a characteristic impairment in this population of patients with MDD (Arnone et al., 2012). Each patient underwent two fMRI scans, one before and one after the manipulation (before the actual start of treatment, i.e., medication/placebo), to capture the neurobiological changes that occur as a result of the manipulation. The study examined whether the effect of the manipulation of the patients’ SL expectancy level on subsequent treatment outcome was mediated by a reduction in amygdala hyperactivation in response to sad faces. The findings supported the proposed mediation model: manipulating SL outcome expectancy was associated with decreased amygdala activation in response to sad faces, which, in turn, was associated with more rapid subsequent reduction in depressive symptoms in the course of antidepressant treatment. If replicated in larger samples, the findings suggest that the effect of SL outcome expectancy manipulation on depressive symptoms is mediated by reduction in amygdala hyperactivation, as measured before patients received antidepressant medication. These findings are consistent with neuroimaging investigations across a range of emotional experiences, from physical pain (Wager et al., 2004) to taste (O’Doherty, Deichmann, Critchley, & Dolan, 2002), suggesting that modulation of amygdala activation is a means by which SL changes in expectancy regulate mood. Thus, it can be suggested that reduction of amygdala hyperactivation in response to sad faces significantly mediates the effect of the manipulation of SL expectancy on subsequent treatment outcome.

The ability to manipulate SL changes in a construct to bring about changes in treatment outcome provides important support for the ability of the construct to function as a mechanism of change (Kraemer, Wilson, Fairburn, & Agras, 2002). Another example of the effect of a manipulation of SL changes on treatment outcome comes from extinction learning during exposure therapy for posttraumatic stress disorder (PTSD). Fear activation is conceptualized as one of the necessary conditions for successful recovery from pathological fear. The other condition is the incorporation of information incompatible with the pathological components of the fear structure (Asnaani, McLean, & Foa, 2016). Repeated therapeutic activation of the trauma memory allows new information to be encoded, reducing the memory-associated fear and anxiety (Foa & Rothbaum, 1998). Extinction learning (behavioral inhibition or fear tolerance) is perceived as an active process involving synaptic modification in the amygdala, and it may be pharmacologically augmented to improve the extinction learning that underlies therapy (Walker, Ressler, Lu, & Davis, 2002). Rothbaum et al. (2014) used D-cycloserine augmentation (vs. placebo) to manipulate extinction learning during exposure therapy for PTSD. Rothbaum et al. (2014) showed that the administration of D-cycloserine resulted in more SL changes in extinction capabilities and in better outcome, as manifested in lower cortisol reactivity and the smallest startle response during virtual reality scenes. Rothbaum et al. (2014) have also shown that hindering or blocking some of the SL changes in fear extension by providing patients...
with alprazolam, which interferes with fear extension, resulted in poorer treatment outcome than achieved with placebo. This example further demonstrates how SL changes can be manipulated to bring about changes in treatment outcome.

### The Tip of the Iceberg: Three Predictions of What Happens When We Make the TLSL Distinction

We have seen only the tip of the iceberg. Much more information is waiting to be discovered about many other mechanisms of change that we expect to be consistent predictors of therapeutic change based on theories and conceptual models, for example, insight (Crites-Christoph et al., 2013) and cognitive change (Lorenzo-Luaces et al., 2015). Until now, none of these mechanisms of change was given a fair chance to show its effect. Based on the presence and direction of the TL and SL effects, it is possible to predict the scenarios under which the effect of mechanisms of change on treatment efficacy is consistent (Table 1). In the first scenario, when both effects (the TL and the SL) are present and act in the same direction, the combined (TL + SL) effect is expected to be consistent. In this case, even without the separation of the TL and SL effects, the combined effect is consistently significant, as in the case of the therapeutic alliance.

In the second scenario, when one of the effects is present and the other is not (either the SL is significant and the TL is not or vice versa), without separating the TL and SL effects the combined effect is expected to be small. Expectancy illustrates this situation. As described above, the meta-analysis (Constantino et al., 2018) shows a relatively small effect without making the TLSL distinction, but the separation shows that only the SL, and not the TL, is significant. Another instance of the second scenario is the association between the therapists’ evaluation of their patients’ understanding and use of core behavioral skills. The TL component of behavioral skills is expected to significantly predict treatment outcome, whereas the TL component may be unrelated to it. Without making the TLSL distinction, the mixed effect may be nonsignificant. With the TLSL distinction, the TL effect is not significantly different from 0, but the SL effect becomes significant (Webb et al., 2019).

In the third scenario, when the two effects act in opposite directions (one is positive, the other negative), the combined effect is expected to be mixed (inconsistent), and separation into TL and SL should produce opposite and significant effects. Examples of the third scenario have been documented in various areas outside of psychotherapy, as, for example, in the relationship mentioned above between exercise and heart attacks. Another example mentioned in the literature is the relationship between typing speed and the number of typos (Hamaker, 2012). Individuals who type faster also have fewer typos. They are simply better typists. This is the TL effect. At the same time, when individuals are requested to increase their typing speed, they are liable to make more mistakes. This is the SL effect. The two effects run in opposite directions. We expect opposite relationships also with respect to insight, as described above and demonstrated in Figure 1. Another example concerns the association between anxiety and depression, where, at least for some individuals, the TL effect is positive and the SL effect negative (Fisher & Boswell, 2016). Whereas the TL effect is consistent with the general view of anxiety and depression as co-occurring, an opposite association emerges for the SL effect: as the levels of depression increase one moment, levels of anxiety are likely to decrease in successive moments.

### Toward Personalized Treatment: Matching Individual TL Components With the SL Targets of Treatment to Produce Patient-Tailored SL Changes

To optimize treatment outcome, we need to identify the TL components that determine whether the basic conditions for treatment are favorable, and the SL changes required to affect treatment outcome for each individual. These TL components can be assembled into a map of the individual’s pathology and strength signature, to serve as treatment moderators (effect modifiers). The SL components, which are the targets of a given treatment, may show greater effect

---

**Table 1**

<table>
<thead>
<tr>
<th>Potential scenario</th>
<th>Case example</th>
<th>TL effect</th>
<th>SL effect</th>
<th>Mixed effect</th>
<th>Consistent combined effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Alliance</td>
<td>+</td>
<td>+</td>
<td>&gt;+</td>
<td>Yes</td>
</tr>
<tr>
<td>2a</td>
<td>Expectancy</td>
<td>ns</td>
<td>+</td>
<td>ns/verysmall</td>
<td>No</td>
</tr>
<tr>
<td>2b</td>
<td></td>
<td>+</td>
<td>ns</td>
<td>ns/verysmall</td>
<td>No</td>
</tr>
<tr>
<td>3a</td>
<td>Insight</td>
<td>–</td>
<td>+</td>
<td>Mixed</td>
<td>No</td>
</tr>
<tr>
<td>3b</td>
<td></td>
<td>+</td>
<td>–</td>
<td>Mixed</td>
<td>No</td>
</tr>
</tbody>
</table>

**Note.** TL = trait-like; SL = state-like; ns = nonsignificant; positive sign (+) = higher levels are related to greater efficacy; negative sign (−) = lower levels are related to greater efficacy.
for some individuals than for others, and for some treatments than for others. Matching the individual’s TL maps with the SL changes that are the targets of each treatment is the key to optimizing treatment efficacy for each individual.

Based on the above definitions, we have the following tools at our disposal to optimize treatment efficacy for each individual (Figure 2): (a) a TL map of the individual’s pathology and strength signature (the colored topography in Figure 2); (b) treatment-specific targets (SL effects) of distinct treatments; and (c) patient-specific mechanisms (SL effects) for distinct subpopulations. To optimize treatment based on the TL map of the individual’s pathologies and strengths (a), we need to find which SL changes that are the target of specific interventions (b) would trigger particular patient-specific mechanisms of change (c).

The TL Maps of Individual Pathology and Strengths

Combining the components of all TL treatment-related constructs of an individual creates a TL map of strengths and deficits. Two patients with the same diagnosis may form different TL maps (e.g., Kotov et al., 2017), with 16,400 different symptom profiles only for MDD (Fried & Nesse, 2015), and coexistence of psychiatric disorders being the rule rather than the exception (Kessler, Chiu, & Walters, 2005). The TL maps can serve to tailor the treatment targets for each patient. For many years, researchers investigating how the patients’ baseline characteristics predict their treatment outcome, focused generally on a single variable at each investigation. They did so whether investigating prescriptive variables (referring to the patient’s ability to benefit from a specific treatment more than another) or prognostic ones (referring to the patient’s ability to show good outcome irrespective of type of treatment; Hollon & Beck, 1986). Advances in the methodological literature, especially the proliferation of the implementations of machine learning approaches in psychotherapy (Cohen & DeRubeis, 2018), make it possible to integrate several baseline variables to create the TL map of the individual’s pathology and strengths signature, presented as a topography rather than as distinct variables. Such maps excel at capturing comorbidities between disorders and at integrating strengths together with deficits, to better reflect the richness and complexity of each individual patient as a human being.

Different SL Effects Underlying Distinct Treatments: Treatment-Specific Targets of Change

Treatment manuals may include different procedures with distinct targets, and are therefore expected to produce distinct SL effects.

Example 1: Cognitive change as a mechanism of change. Conceptually, in cognitive treatment, the use of cognitive techniques is expected to bring about symptomatic reduction by changing negatively biased beliefs and thinking styles (Beck, Rush, Shaw, & Emery, 1979). Empirical findings, however, failed to show consistent support for the specificity of this effect, which was reported in other types of psychotherapy as well (Garratt, Ingram, Rand, & Savalani, 2007; Warmerdam, van Straten, Jongsma, Twisk, & Cuijpers, 2010), and even in antidepressant treatment (Segal et al., 2006, but see also DeRubeis et al., 1990). It is possible that the TL effect is common across various treatment modalities, so that patients reporting more negatively biased thinking may tend to have better prognosis, perhaps because they have greater insight into their depression. By contrast, it is reasonable to expect that the SL effect of cognitive change is a specific target of cognitive treatment. Support for the potential importance of the TLSL distinction in revealing the specificity of cognitive change is found in a recent study that focused on the SL effect of cognitive change in non-cognitive–behavioral therapy (non-CBT) treatments. The study shows that when focusing on the SL effect, cognitive change was no longer a significant predictor of outcome in non-CBT therapy, such as psychodynamic and antidepressant treatment (Zilcha-Mano, Chui, et al., 2016). The study illustrates the erroneous conclusion that may be reached without making the TLSL distinction, that cognitive change has no specificity, but rather is a mechanism of change in any type of treatment, psychotherapy or pharmacotherapy.
Different SL Effects for Distinct Subpopulations of Patients: Patient-Specific Mechanisms of Change

Distinct subpopulations of patients may show different responses to given SL changes. This type of information can tell the therapist which mechanisms of change are most effective in bringing about change for a given individual, and should therefore be the focus of treatment.

Example 1: cognitive changes. A recent study by Fitzpatrick, Whelen, Falkenström, and Strunk (2020) showed that in cognitive therapy for depression, SL cognitive changes (e.g., adopting a more realistic view) predicted improvement in symptoms more robustly for patients with fewer perceived social skills and for those with greater interpersonal problems.

Example 2: Gaining insight from treatment sessions. A study by Bounoua et al. (2018) suggests that the SL effect of insight gained in a session by adolescents may depend on their emotion regulation capacities. Poorer emotion regulation capacities at baseline increased adolescents’ vulnerabilities to spillover effects, in which negative interpersonal events from the past week influence pre-session negative affect and spill over to the adolescents’ ability to gain insight from their treatment sessions.

Example 3: Working alliance. The question of who are the individuals for whom SL changes have the greatest influence on treatment outcome received the most attention in research on the working alliance (Constantino et al., 2017; Lorenzo-Luaces, DeRubeis, & Webb, 2014). Without the TLSL distinction, the alliance seemed to contribute about equally to treatment success for everyone, in some 295 independent samples and more than 30,000 patients (Flückiger et al., 2018). But the TLSL distinction reveals new information that matches much of the experience accumulated over 100 years of clinical work: the SL component of alliance plays a more therapeutic role for patients who have problems forming satisfying relationships with others outside the treatment room, especially individuals with personality problems (Falkenström, Granström, & Holmqvist, 2013) and those suffering from low interpersonal agency who report problems with submissiveness (Penedo et al., 2019). For example, clustering analyses of the alliance assessments of patients at the first four sessions of treatment reveals that some patients show more SL change than others. For patients with more interpersonal problems, more profound SL changes were associated with greater subsequent treatment outcome than for patients with fewer interpersonal problems. This finding has been hidden for many years under data that mixed together the TL and SL components of alliance, and it demonstrates that a given subpopulation with specific interpersonal characteristics benefits more from the strengthening of the alliance for achieving better treatment outcome.

Optimizing Treatment by Matching TL Maps With Treatment-Specific Targets and Patient-Specific Mechanisms of Change

Distinct SL changes may serve as mediators underlying the effects of certain treatments. Thus, they may answer the question how treatment works by pointing to the processes that are targeted by treatment procedures and techniques to bring about therapeutic change (Kazdin, 2007, 2009). The individuals’ TL signature map may serve as a group of moderators, identifying the patients for whom a given treatment may work best. Combining the two questions that have been at the heart of psychotherapy research for years, how treatment works and for whom, has the potential to reveal the distinct processes involved in the successful treatment of individual patients, targeting patient-specific mechanisms of change to optimize treatment outcome (Kazdin, 2007; Zilcha-Mano, 2018). Moderated mediation statistical models were designed to answer the combination of the two
questions into one question: “How does treatment work for a particular individual?” It has been argued that moderated mediation models are instrumental for advancing toward personalized psychotherapy (Cuijpers et al., 2019; Hollon, 2019).

The TLSL distinction provides concrete instructions on how to identify the moderator and the mediator that can be combined together to create a moderated mediation model. Both the homeostasis and the strength approaches can guide the use of the TLSL distinction in testing moderated mediation models aimed at spotlighting the patient-specific mechanisms of change by identifying the SL changes that are required to optimize treatment outcome for a given patient according to the individual’s TL map.

According to the homeostasis approach, the TL component defines what “went out of order,” that is, the individual’s signature pathology, which, in turn, reveals what is needed to return to homeostasis. Treatment will result in better outcomes if it targets the relative deficits of the patients, helping them acquire skills and capacities they do not yet possess. An example of how TL deficits can be used to determine the SL changes that are needed based on the homeostasis approach comes from a study by Barber and Muenz (1996), demonstrating that individuals with higher levels of avoidant personality may benefit most from cognitive therapy, whereas individuals with higher levels of obsessive personality may benefit most from interpersonal treatment. The authors noted that their findings are consistent with the “theory of opposites,” meaning that helpful treatment requires therapists to provide patients with opportunities to participate in interpersonal behaviors that go against their general tendencies.

According to the strength approach, the target of a therapeutic intervention should be the patients’ TL signature strengths, coopting the patients’ most adaptive capabilities of perceiving and acting in the world (Cheavens, Strunk, Lazarus, & Goldstein, 2012). An example of how TL strengths can be used to determine the SL changes that are needed based on the strength approach comes from a study by Cheavens et al. (2012), in which patients were randomized to conditions where they were treated using cognitive–behavioral strategies that target either their relative strengths or the opposites of these strengths, that is, their deficits. Findings suggest that personalizing treatment to patients’ relative strengths led to better outcomes than did treatment personalized to their relative deficits.

Thus, whether aiming at homeostasis, at capitalizing on strength, or both, the TL \( \times \) SL interactive effect is the key to targeting patient-specific mechanisms of change. Without the TLSL distinction, the interaction between them is not even conceivable. Paraphrasing Gordon Paul’s (1967) iconic question, part of what we may be able to answer using the TLSL distinction is which core SL changes should be the target of intervention with a particular patient, given this patient’s TL signature map? We may be able to at least partially answer this question with the help of the TLSL distinction. The integration of the TLSL distinction with moderated mediation models focuses the investigation on patient-specific mechanisms of change, so that the TL signature map of the individual across key constructs determines the SL changes that are most needed to optimize treatment outcome. When the goal of treatment is reaching homeostasis, one needs to identify the constructs in which TL deviation from homeostasis is apparent, then focus on producing SL changes in that particular construct. For example, in exposure therapy for irritable bowel syndrome, SL reductions in behavioral avoidance were associated with better treatment outcome. This was especially the case for individuals with higher TL avoidance (Hesser, Hedman-Lagerlöf, Andersson, Lindfors, & Ljótsson, 2018). Here again, the TL and SL effects run in opposite directions, and their interaction is pivotal for progress toward personalized treatment. When the goal of treatment is capitalizing on strengths to empower the patient, one needs to identify the specific constructs that serve as the individual’s TL strengths, then focus on producing SL changes in that particular construct. Naturally, the two goals can be integrated. Identifying the SL changes essential for a given individual makes it possible to design the most accurate treatment plan for that individual.

**Summary**

Although the TLSL distinction is not a solution for all mechanism-based questions, and naturally, not to all psychotherapy research questions, it represents an important piece of the complex puzzle that is the black box of psychotherapy. The puzzle contains many other pieces, including therapists’ TL characteristics, training, and development (Baldwin & Imel, 2013), responsiveness (Hatcher, 2015; Stiles, 2013) not only to the patients’ TL characteristics but also to ever-changing SL components, multicultural issues (Owen et al., 2016), idiographic data of the individual patient (Fisher et al., 2018; Silberschatz, 2017), practice-oriented research (Castonguay & Muran, 2015), integration of distinct methods of inquiry (Hill, 2012), interdisciplinary approaches to measurement (Zilcha-Mano & Ramseyer, 2020), and identifying cross-theoretical principles of change (Goldfried, 2019), among others. The TLSL distinction itself can benefit from knowledge accumulating in other fields of science outside of psychotherapy research, such as advances in the methods available for disentangling within- and between-individuals variances.

The TLSL distinction is expected to enhance the consistency of results concerning mechanisms of change, and harnessing the TL \( \times \) SL interactive effect can pave the road to optimizing treatment efficacy. It is not necessary to put thousands of patients through controlled studies again. The
knowledge we need is already cached in legacy data. All that is needed is to revisit studies that have examined the mechanisms of change several times over the course of treatment, make the TLSL distinction, and test the TL × SL interactive effects. Retrospective analysis of already collected data should be complemented by studies designed a priori to investigate the TL and SL effects of each construct of interest on treatment efficacy. Such studies will be able to implement the required design, with the appropriate assessment schedule for each construct (pretreatment, in-treatment, and posttreatment; see the online supplemental materials) to extract information about both the TL characteristics of individual patients and the process of therapeutic change. Once the black box of psychotherapy lies open, an important step in the progress toward evidence-based, personalized treatment can be taken.

Clinical Demonstration

At the start of treatment, the therapist can systematically evaluate the individual’s signature TL components. The evaluation can be accomplished using feedback systems (Lutz, Rubel, Schwartz, Schilling, & Deisenhofer, 2019), clinical interviews (Hoffman, 2020), a diagnostic battery, or any other assessment approach. The result is an initial formulation (topography) of the individual’s main TL strengths and weaknesses that serves as a map for identifying the active ingredients of treatment most critical for that individual to achieve SL therapeutic change. For example, a patient may arrive at treatment with a TL signature that includes high potential to form a strong alliance (measured by the alliance expectation questionnaire; Barber et al., 2014), a moderately high level of insight (measured by the SUIP-I; Gibbons & Crits-Christoph, 2017), and a low level of emotional arousal (measured by the Client Emotional Arousal Scale–III (CEAS-III); Carryer & Greenberg, 2010; Warwar & Greenberg, 1999). Based on this TL topography, the therapist must choose from a repertoire of evidence-supported therapeutic techniques the ones that are likely to be most effective with the individual patient. Candidates may be, for example, identifying and repairing alliance ruptures (Safran & Muran, 2000), insight-based work (Høstglen et. al., 2006; Luborsky, 1984), and emotion-focused techniques (Elliott, Watson, Goldman, & Greenberg, 2003).

The patient’s TL components are likely to be related to each other. Difficulty in emotional awareness and expression may affect and be affected by one’s level of insight. Difficulty in emotional awareness and expression is also affected by and affecting the ability to form satisfying intimate relationships with others, and may affect the ability to form an intimate relationship with the therapist. Therefore, the individual in question may benefit from all three types of therapeutic techniques mentioned above, and from others, including cognitive and behavioral techniques.

Yet, given that time in treatment is limited and that one may have to choose one’s battles (or at least one’s port of entry), selecting the most effective techniques for any individual patient from the many available is likely to be more cost-effective than randomly selecting one of them (Cohen & DeRubeis, 2018). In a clinical decision-making process of this type, it is therefore of great interest to identify the active ingredients expected to bring the most consequential SL change for that individual, based on the individual’s TL signature. In the above example, the patient is likely to benefit most from techniques focused on the patient’s most impaired TL characteristics, in this case, the ability to develop emotional awareness, acceptance, and regulation skills. Emotion-focused (Elliott, Watson, Goldman, & Greenberg, 2003) and acceptance and commitment techniques (ACT; Hayes, Strosahl, & Wilson, 2011) are good candidates. The therapist and patient can build on the patient’s TL insight abilities and the capability of forming strong relationships. The capacity to form a strong alliance can be instrumental in achieving a genuinely shared engagement in a safe, supporting therapeutic relationship, with an empathically attuned and responsive therapist. Such TL ability may enable the patient to express and explore emotions, knowing that a safe haven is present in times of need. The patient’s moderately high TL level of insight may be instrumental in developing the motivation needed for the strenuous emotional work of accessing, experiencing and expressing painful emotions, based on an understanding of the need for more flexible management of emotions for dealing with the main difficulties the patient faces. Such emotion-focused therapeutic work is likely to result in gains in alliance, cognitive and behavior changes, and greater insight, perhaps not as the driving forces of change but rather as products of effective treatment in this individual case. If time in treatment allows, choosing a cocktail of techniques targeting distinct TL deficits of the individual may be advisable, to be implemented in an integrative (Wachtel, 2014) or a modular manner (Barlow et al., 2017), starting with the TL component that has the most adverse effect on the patient’s life. Highly effective treatment may even result in changes in the patient’s TL characteristics.

References


Received September 16, 2019
Revision received February 11, 2020
Accepted February 26, 2020