

# Conceptualizing Epistemic Trust in Psychotherapy: A Triadic Model

Shimrit Fisher, Sigal-Zilcha-Mano, Peter Fonagy

## Introduction

In today's rapidly changing social environment, people face the challenge of determining whom they can safely trust and who will lead them astray. Dealing with this challenge is crucial not only for individuals' survival but also for their adaptation to social norms, habits, and the unstated rules of culture. However, people do not navigate the ever-changing social world blindly; rather, Nature has endowed humans with the ability to be influenced by new information under certain circumstances and wary of it in others. This evolutionary ability to correctly identify others as trustworthy and hence rely on the information they convey as personally relevant and generalizable is known as Epistemic Trust (ET; Fonagy & Campbell, 2017) and enables individuals to benefit from positive social communications and avoid deception. Theoretical conceptualizations suggest that ET develops at a young age within attachment relationships (Fonagy and Allison., 2014). Infants are instinctively look for cues from their primary caregivers that will allow them to better understand and act in ambiguous situations (Corriveau et al., 2009; Fonagy et al., 2017; Ronfard & Lane, 2019). Positive cues such as eye contact, turn-taking contingent discourse, and being called by name indicate to the infant that new and relevant information is being conveyed, which is thus worthy of being retained (Fonagy et al., 2017). Growing up in a sensitive caregiving environment equips children with the ability to better identify potentially misleading information, thus allowing them to be justifiably vigilant (Fonagy et al., 2019). In contrast, under predominantly hostile caregiving or insecure attachment relationships within the infant's early environment, the ability to discriminate reliable from untrustworthy communication may be severely hampered (Campbell et al., 2021). These individuals may become excessively cautious about learning from others and discriminate poorly between genuinely trustworthy and potentially false social information. As a result, their ability to update their understanding of social situations is curtailed and they appear inflexible in the face of rapid social change (Fonagy et al., 2019).

Depending on their distinct developmental trajectories, individuals will develop relatively stable trait-like characteristics of ET that may dictate different types of functioning in social and interpersonal contexts as adults. Some individuals may adjust their inner position safely in light of new information or experiences (Fonagy et al., 2018), may rely on their agentic sense of self and remain open to learning about themselves and others. Other individuals rely on the same fixed models of relating to the self and others, which prevents them from accepting new valuable information (Luyten et al., 2019). This perspective assumes that by definition, part of the construct of ET is considered to be a stable, quasi trait-like capacity which is formed during childhood and serves as a template in adult relationships. Nevertheless, ET is also subject to formative social influences throughout the lifespan. A range of interpersonal contexts can offer differential exposure to reliable sources of information in which social learning can occur. Teachers, peers, social media, and psychotherapy may change people's general expectations of trustworthiness (Luyten et al., 2020).

## ET in psychotherapy: the triadic model

Psychotherapy can draw on patients' epistemic trust as leverage to safely challenge inner working models and enable them to acquire the skills needed to navigate the social environment (Fonagy

et al., 2019). However, psychotherapy can also serve as an essential context for creating state-like changes in ET itself. When the ET trait-like component is well-established, genuine curiosity can develop about the self and others, which fosters the ability to learn and deduce from one interpersonal context to another. However, in cases where ET is not sufficiently established, its restoration can become the core driver of effective therapeutic change (Fonagy et al., 2015; Luyten et al., 2019). Identifying a person's ET levels and, if necessary, using ET as a mechanism of change requires understanding how it is expressed and its components in interpersonal relationships in general and in treatment in particular. Here, three components are suggested to capture the construct of ET in psychotherapy: sharing, the we-mode, and learning (Fisher et al., 2020).

**Sharing** refers to a person's willingness to discuss meaningful experiences and the thoughts and feelings that ensue. Communicating meaningful experiences can create a sense of companionship or allyship, which can mitigate a person's feelings of isolation. However, it also contributes to fulfilling various other possible psychological functions such as clarifying vague emotions, creating a logical order in one's experiences, broadening one's perspectives of the social world, etc. In psychotherapy, patients are invited to share; that is, to present thoughts or experiences so that another person (the therapist) can envision the same event and consider its effects (Fonagy & Campbell, 2017). The willingness to share experiences in therapy is assumed to reflect patients' trait-like approach toward relationships and their long-term views on others (including the therapist) in their social network (Fonagy et al., 2019). When engaging in therapy, some patients are better able to 'tune into' their social environment and perceive the therapist's intentions as beneficial or at least benign, which prompts them to believe that sharing their conflicts is worthwhile. These patients may show curiosity regarding the therapist's input and can enjoy a fruitful dialogue within the therapeutic session. Other patients, however, may be ambivalent about the value of sharing. Because they view the social world as generally malevolent, these patients covertly expect to be misunderstood, be subject to criticism, and made to feel guilt or shame, which may lead them, at least partially, to abandon the possibility of sharing altogether. A shift towards greater willingness to share in these patients is thus crucially important since sharing heightens the possibility that their discourse will be accurately understood and appreciated. The willingness to share within therapy constitutes the foundation of all therapeutic processes regardless of specific therapeutic modalities. It also mediates another essential component of this process: the willingness to share with others outside the therapy room. Sharing outside therapy is a key part of the change process since interpersonal relationships can be reconfigured and scaffold important adjustments in self-perception. Improving the capacity and motivation to share self-experiences with others is likely to contribute to the individual's well-being and adaptation to life's challenges by reinforcing intra- and inter-personal changes.

**"We-mode"** When an individual is open to sharing experiences, thoughts, and feelings with another, a co-creation of mutual communication is enabled. The delicate moments in which two conversation partners acknowledge each other's perspectives and reflect on the same piece of subjective reality are often termed the "we-mode" (Choi-Kain et al., 2022). In we-mode moments, partners establish joint attention to explore and better understand one another's emotional perspectives (Bo et al., 2017). Therefore, we-mode moments can help assuage a sense of loneliness and alienation (Fonagy et al., 2019). Engaging in we-mode interactions enhances the ability to

understand oneself and others and helps provide the individual with new ways of acting and reacting to others (Gallotti & Frith, 2013).

At its most fundamental, psychotherapy works by exchanging and processing information about oneself and others. Irrespective of the therapeutic approach, it constitutes a setting where minds can meet; i.e., establish we-mode moments (Bo et al., 2017). Despite the central role of the we-mode in all types of therapy, the patient-therapist dyadic ability to achieve and maintain these moments can vary widely. Some patients engage in frequent we-mode moments more easily so that therapeutic interventions can be grounded on this personality facet to achieve change. Other patients may find it almost impossible to feel that their "personal truth" can be recognized and conveyed (Fonagy et al., 2019). When working with these patients, the therapeutic goal needs to focus on achieving, establishing, and preserving such moments. Either way, the patient and therapist must mutually explore and acknowledge what happens in the patient's mind when interacting with social partners, including with the therapist (Fonagy et al., 2019). This sense-making process gives conscious meaning to inner narratives and may allow for an internal dialogue to take place in the patient's mind (Benjamin, 2004) and with others (Fonagy et al., 2019; Luyten et al., 2020).

**Learning** is the process of acquiring new knowledge about how the social environment operates, which can then be applied to other life situations (Csibra & Gergely, 2006; Reed et al., 2010). When navigating the social world, individuals need to recognize when and with whom it is safe to open their minds to acquiring new knowledge, and under what circumstances it is not advisable (Sperber, 2001). Being open and sensitive to new knowledge and experiences can be an advantage as long as individuals can adapt flexibly to their ever-changing environment. This is because the social environment is constantly changing, thus providing individuals with countless opportunities for absorbing and integrating social knowledge. Therefore, an individual must be able to learn about opaque cultural artifacts or interpersonal skills which do not reveal their meaning or function upon simple observation. The mechanism that allows this information to be transmitted is activated through human communication (Csibra & Gergely, 2006; 2009; 2011). For the safe activation of this mechanism, nature endowed humans (learners) with the capacity to attend to others to receive key information essential to survival (Csibra & Gergely, 2011). Human beings have preserved this unique sensitivity to signs indicating that information is intentionally directed at them and, therefore, is to be tracked. (Fonagy et al., 2017). We have argued that achieving a meeting of the minds in situations such as the we-mode may be a sign or signal, regardless of age, which generates the willingness to learn from a social situation.

Psychotherapy can be considered a learning opportunity since it is built upon humans' innate evolutionary capacity for social learning. It provides patients with a safe and reassuring context to be exposed to new knowledge they can then apply to other interpersonal contexts. However, learning in therapy depends on the patient's degree of openness to new information and events that unfold during therapy (Fonagy et al., 2015). Patients capable of social information exchange ("good" learners) can easily update their knowledge about themselves and others during therapy. Other patients may fail to benefit from benign social interactions, including in therapy ("poor" learners). In order to revive the ability to learn, they must first be able to calm their a-priori apprehensiveness towards new information and informants. Whether learning from social situations is the purpose or the product, by enabling patients to revisit their perceptions of themselves and the world, psychotherapy can pave the way towards further learning outside therapy (Fonagy et al., 2017).

## **Communicating Vessels**

All three components are theoretically essential parts of ET (Fisher et al., 2020). When individuals put their overall experience into perspective and calibrate their mind to those of others (sharing) while establishing a mutual discourse for the processing of ideas (we-mode), new information pertinent to social adaptation (learning) can be acquired. However, these three postulated components can manifest differently within psychotherapy, depending on the patient's trait-like ET characteristics. Nevertheless, the process of psychotherapy is one of the critical factors that can bring about state-like changes in ET by making the patient open to social learning and personal change (Luyten et al., 2021). Since psychotherapy is a form of social communication that provides the opportunity to learn, achieving state-like changes in ET during therapy can be a valid treatment goal insofar as a change in ET makes the broader change process through social learning possible. One potential mechanism to bring about state-like changes in ET is associated with therapists' recognition and articulation of their patients' more fine-grained self-experiences (Fonagy et al., 2019). Feeling accurately reflected in the therapist's mind may pave the way to social learning and the restoration of ET in the patient (Fonagy & Target, 2006; Sharp et al., 2020). Figure 1 illustrates how feeling reflected accurately and validated appropriately (we-mode) can allow patients to be more permeable to new information (learning) that can modify internal working models (sharing). The therapist's explicit effort to see the world from the patient's standpoint (i.e., person-centered intervention) may be instrumental in facilitating this process of change by making the patient feel safe enough to share openly and learn (Fisher et al., 2020).

[INSERT FIGURE HERE]

Figure 1: Theoretical Model of trait-like characteristics and state-like changes in ET in psychotherapy. Establishing the 'we-mode' in therapy enables learning about the self, others, and the viability of sharing.

## **Present and Future Directions**

ET theory considers acquiring new information through social interactions as essential to people's sense of security, confidence, and survival (Bo et al., 2017; Csibra & Gergely, 2009; Fonagy et al., 2015). In recent years, a number of theoretical works have explored the potential implications of ET for psychotherapy (Fonagy et al., 2017; 2019). Clinicians and researchers have shown a growing interest in understanding the contribution of ET to successful treatment outcomes (e.g., Bo et al., 2017). However, little is known empirically, particularly in terms of studying the contribution of each component (sharing, we-mode, and learning) separately. To date, there is no method or measure to quantify patients' level of ET within a therapeutic relationship (Folmo et al., 2019). Empirical assessments could provide an effective framework to characterize individual patients and their outcomes and the ways in which the development of ET unfolds in therapy. Focusing on individual differences can help identify how psychotherapy works for certain subpopulations (Zilcha-Mano, 2021) since certain subpopulations with various difficulties are likely to benefit differently from therapeutic interventions, depending on their trait-like ET. Future studies should thus focus on whether specific interventions by the therapist can bring about state-like changes in ET and whether this, in turn, can translate into beneficial treatment outcomes. Overall, focusing on patients' ET can enable clinicians to tailor the therapeutic approaches used in treatment for each patient's particular needs.

## References

- Bo, S., Sharp, C., Fonagy, P., & Kongerslev, M. (2017). Hypermentalizing, attachment, and epistemic trust in adolescent BPD: Clinical illustrations. *Personality Disorders: Theory, Research, and Treatment*, 8(2), 172.
- Benjamin, J. (2004). Beyond doer and done to: An intersubjective view of thirdness. *The Psychoanalytic Quarterly*, 73(1), 5-46.
- Campbell, Chloe, Michal Tanzer, Rob Saunders, Thomas Booker, Elizabeth Allison, Elizabeth Li, Claire O'Dowda, Patrick Luyten, and Peter Fonagy. "Development and validation of a self-report measure of epistemic trust." *PLoS one* 16, no. 4 (2021): e0250264.
- Choi-Kain, L. W., Simonsen, S., & Euler, S. (2022). A Mentalizing Approach for Narcissistic Personality Disorder: Moving From "Me-Mode" to "We-Mode". *American Journal of Psychotherapy*, appi-psychotherapy.
- Corriveau, K. and Harris, P.L. and Meins, E., Fernyhough, C., Arnott, B., Elliott, L., Liddle, B., Hearn, A., Vittorini, L. & De Rosnay, M. (2009). Young children's trust in their mother's claims: longitudinal links with attachment security in infancy. *Child Development*, 80 (3). 750-761.
- Csibra, G., & Gergely, G. (2006). Social learning and social cognition: The case for pedagogy. *Processes of change in brain and cognitive development. Attention and performance XXI*, 21, 249-274.
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in cognitive sciences*, 13(4), 148-153.
- Csibra, G., & Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 366(1567), 1149-1157.
- Egyed, K., Király, I., & Gergely, G. (2013). Communicating Shared Knowledge in Infancy. *Psychological Science*, 24(7), 1348–1353.
- Fisher, S., Guralnik, T., Fonagy, P., & Zilcha-Mano, S. (2021). Let's face it: video conferencing psychotherapy requires the extensive use of ostensive cues. *Counselling Psychology Quarterly*, 34(3-4), 508-524.
- Folmo, E. J., Karterud, S. W., Kongerslev, M. T., Kvarstein, E. H., & Stänicke, E. (2019). Battles of the comfort zone: modeling therapeutic strategy, alliance, and epistemic trust—a qualitative study of mentalization-based therapy for borderline personality disorder. *Journal of Contemporary Psychotherapy*, 49(3), 141-151.
- Fonagy, P., & Allison, E. (2014). The role of mentalizing and epistemic trust in the therapeutic relationship. *Psychotherapy*, 51(3), 372–380.
- Fonagy, P., & Campbell, C. J. P. H. (2017). Mentalizing, attachment and epistemic trust: how psychotherapy can promote resilience. *Psychiatria Hungarica*, 32(3), 283-287.
- Fonagy, P., Campbell, C., & Bateman, A. (2017). Mentalizing, attachment, and epistemic trust in group therapy. *International Journal of Group Psychotherapy*, 67(2), 176-201.

- Fonagy, P., Gergely, G., & Jurist, E. L. (Eds.). (2018). *Affect regulation, mentalization and the development of the self*. Routledge.
- Fonagy, P., Luyten, P., & Allison, E. (2015). Epistemic petrification and the restoration of epistemic trust: A new conceptualization of borderline personality disorder and its psychosocial treatment. *Journal of personality disorders, 29*(5), 575-609.
- Fonagy, P., Luyten, P., Allison, E., & Campbell, C. (2019). Mentalizing, epistemic trust and the phenomenology of psychotherapy. *Psychopathology, 52*(2), 94-103.
- Fonagy, P., & Target, M. (2006). The mentalization-focused approach to self pathology. *Journal of personality disorders, 20*(6), 544-576.
- Gallotti, M., & Frith, C. D. (2013). Social cognition in the we-mode. *Trends in cognitive sciences, 17*(4), 160-165.
- Luyten, P., Campbell, C., Allison, E., & Fonagy, P. (2020). The mentalizing approach to psychopathology: State of the art and future directions. *Annual review of clinical psychology, 16*, 297-325.
- Luyten, P., Campbell, C., & Fonagy, P. (2019). Reflections on the contributions of Sidney J. Blatt: The dialectical needs for autonomy, relatedness, and the emergence of epistemic trust. *Psychoanalytic Psychology, 36*(4), 328.
- Luyten, P., Campbell, C., & Fonagy, P. (2021). Rethinking the relationship between attachment and personality disorder. *Current Opinion in Psychology, 37*, 109-113.
- Reed, M. S., Evely, A. C., Cundill, G., Fazey, I., Glass, J., Laing, A., ... & Stringer, L. C. (2010). What is social learning?. *Ecology and Society, 15*(4).
- Ronfard, S., & Lane, J. D. (2019). Children's and adults' epistemic trust in and impressions of inaccurate informants. *Journal of Experimental Child Psychology, 188*, 104662.
- Sharp, C., Shohet, C., Givon, D., Penner, F., Marais, L., & Fonagy, P. (2020). Learning to mentalize: A mediational approach for caregivers and therapists. *Clinical Psychology: Science and Practice, 27*(3), e12334.
- Sperber, D. (2001). An evolutionary perspective on testimony argumentation. *Philosophical topics, 29*(1/2), 401-413.
- Zilcha-Mano, S. (2021). Toward personalized psychotherapy: The importance of the trait-like/state-like distinction for understanding therapeutic change. *American Psychologist, 76*(3), 516.