Mentalizing, epistemic trust and the phenomenology of psychotherapy

Peter Fonagy[1*], Patrick Luyten[1,2], Elizabeth Allison[1], Chloe Campbell[1]

[1] Research Department of Clinical, Educational and Health Psychology, University College London, UK

[2] Faculty of Psychology and Educational Science, KU Leuven, Leuven, Belgium

*Corresponding author

Professor Peter Fonagy

Research Department of Clinical, Educational and Health Psychology

University College London, UK

Tel: 0044 7679 1474

Email: p.fonagy@ucl.ac.uk

Keywords

Psychotherapy, phenomenology, mentalizing, epistemic trust

Abstract

This paper seeks to elucidate the phenomenological experience of psychotherapy in the context of the theory of mentalizing and epistemic trust. We describe two related phenomenological experiences that are the domain of psychotherapeutic work. The first is the patient's direct experience of their own personal narrative being recognised, marked and reflected back to them by the therapist. Secondly, this intersubjective recognition makes possible the regulation and alignment of the patient's imaginative capacity in relation to our phenomenological experiences. In describing three aspects of the communication process that unfold in effective psychotherapeutic interventions – 1) the epistemic match, 2) improving mentalizing, and 3) the re-emergence of social learning – the way in which any effective treatment is embedded in metacognitive processes about the self in relation to perceptual social reality is explained. In particular, attention is drawn to wider social determinants of psychopathology. We discuss the possible mechanism for the relationship between the social-economic environment and psychopathology, and the implications of this for psychotherapeutic treatment.

Introduction

Historically, although there were moments of exchange between phenomenological approaches and psychoanalysis, there has also been something of a divide, with the two approaches developing in relative isolation [1]. This mainly has to do with the traditional emphasis in psychoanalysis on the role and nature of unconscious processes, whereas the phenomenological tradition typically focusses on the here and now of conscious, subjective experience. We believe that this has been unfortunate because any effective approach in psychotherapy should not only focus on the underlying psychological processes involved in the human mind, but also the more experience-near subjective functioning of individuals. This will be the central focus of this paper.

In psychoanalytic thinking, consciousness has been regarded as the less significant hand tool of the real work of mental activity which takes place unconsciously. In this respect, psychoanalytic thinking – never normally chary about broaching the big questions of human experience – has not directly engaged with the "hard problem of

consciousness" [2]. In this paper we argue that contemporary psychoanalytic thinking about mentalizing and epistemic trust provides us with a way into thinking about human consciousness and the phenomenological experience of psychotherapeutic intervention from which its effectiveness derives. Mentalizing is the term used to describe a particular facet of the human imagination: an awareness of mental states in oneself and in other people, particularly in explaining their actions. Epistemic trust is defined as openness to the reception of social knowledge that is regarded as personally relevant and of generalizable significance [3]. We will thus try to demonstrate that the process of *jointly* undertaking the metacognitive work of considering the subjective experience of being in possession of a mind is at the heart of the psychotherapeutic project [4].

Psychotherapy is centrally concerned with the problem of subjectivity, a shadowy artefact of human consciousness that is nevertheless a necessary construct for psychological health. Recent developments in thinking about the intrinsically social nature of higher order cognitive function would lead us to add another critical adaptive purpose to the evolution of subjectivity: helping us to manage complex social relationships. This suggestion might be understood as part of a broader shift in thinking about the origins and functions of some of the characteristics that we identify as central to our identity as a species as serving the transmission of culture and social capabilities. A major expression of this is found in Mercier and Sperber's book *The Enigma of Reason*, in which they examine the question of the human capacity for reason, alongside an equally pronounced capacity for irrationality [5]. They argue that reason is primarily social, that the function of logic and reason is to enable us to cooperate, negotiate and agree social terms with others. A recent hypothesis presented by Mahr and Csibra argued, similarly, that one of the functions of episodic memory is to enable social communication [6]. Thus, memory of personal experience ensures that we have justifications for why we believe what we do, for keeping track of where we are placed in terms of obligations and commitments with others, whom we can rely on and whom we should regard with caution. These are key elements that enable the social cooperation, social learning and the construction of a network of relationships that make culture and its transmission possible [6]. We have argued, in a similar vein, that human consciousness evolved to allow us to share our experiences, to communicate a "shared narrative" on which relationships, social ties, and group cohesion can all be built [7]. We become conscious of those aspects of the world which others reflect on as well, and this applies equally to our internal subjective world [7].

Mentalizing, meta-cognition and social communication

This paper will discuss the phenomenology of psychotherapy in the light of this thinking about the inter-relational drivers of the human experience of subjectivity, and we will suggest that the effectiveness of treatment derives from the experience of social metacognition, which triggers a capacity for social learning. Mentalizing theory, to put it in phenomenological terms, is explicitly based around the human experience of qualia, mainly but not exclusively social qualia, and in particular metacognition in relation to qualia. Qualia refers to the perception of the quality of an experience – mentalizing is about the perception and interpretation of behaviour and thus from a phenomenological perspective concerns the process of reflection in relation to qualia. As with subjectivity, the existence of qualia has been disputed by philosophers of the mind, but whether or not qualia exist [8], there can be no debate that our *perception* of qualia exists, and that the loss or distortion of that experience is psychically highly disruptive and potentially terrifying. There is consistent evidence that any form of mental disorder is associated with a temporary or chronic distortion of internal and/or external reality [9]. Awareness in relation to mental states is perforce bound up with the emergence of imagination. Beyond a biological resonance to others' emotional states, we have to *imagine* their phenomenological experience, their thoughts and feelings. Indeed, the leap forward in the development of the complexity of human tools requiring collaboration in construction coincides with the emergence of human objects which are the products of imagination, e.g., the torso of lion and the head of a human, and complex representations (such as cave drawings) [10, 11]. With imagination, of course comes the potential of unhelpful, indeed, frightening possibilities.

In individuals who are seriously disrupted in their capacity to mentalize, qualia become disrupted – in the first instance, this is most conspicuous in relation to social and emotional experiences. But it is not limited to such. Commonly cited examples of qualia are subjective experiences of colour or the smells of food – distress or distortions in thinking are all capable of disrupting the ways in which we might enjoy, interpret or react to such perceptual experiences. Martin Debbané's work on psychosis continuum and the role of early, subtle mentalizing disruptions which may precede more conspicuously aberrant mentalizing difficulties is relevant here [12]. Descriptions of the non-mentalizing modes of psychic equivalence, pretend mode and the teleological mode all

involve, according to the individual, distortions in the way in which he or she responds to perceptual experiences of the real world. In the psychic equivalence mode, thoughts and feelings become "too real" to a point where it is extremely difficult for the individual to entertain possible alternative perspectives. When mentalizing gives way to psychic equivalence, what is thought is experienced as being real and true, leading to what clinicians describe as "concreteness of thought" in their patients. In the teleological mode, states of mind are recognized and believed only if their outcomes are physically observable. Hence, the individual can recognize the existence and potential importance of states of mind, but this recognition is limited to very concrete situations. For example, affection is perceived to be true only if it is accompanied by physical contact such as a touch or caress. In the pretend mode, thoughts and feelings become disconnected from reality. In more extreme cases, this may lead to feelings of derealization and dissociation.

We have previously suggested that, for example in working with an adolescent with emerging borderline traits, that there might be value in an initial approach being a physical one, such as running with them, and discussing with them what the experience of running was like. This thinking is based on the idea that individuals who are really poor at mentalizing require not just cognitive interventions, but ones that relate to the body more directly: it is not possible to access mentalizing if the self is overwhelmed by negative interference which impairs normal cognitive function. As mentioned, Debbané's research on mentalizing and psychosis has extended this line of thinking [12, 13], as well as Fotopoulou and Tsakiris's work on embodied mentalizing [14].

Developing this thinking, we would like to describe here two related phenomenological experiences that occur in psychotherapeutic work. The first is the patient's direct experience of their own personal narrative being recognised, marked and reflected back to them by the therapist. Secondly, this intersubjective recognition makes possible the regulation and alignment of our imaginative capacity in relation to our phenomenological experiences. The harnessing of these two metacognitive activities – a) the conscious recognition of one's personal narrative by another (with the intersubjective acknowledgment of subjectivity that this brings) and b) assistance in regulating the social imagination – generate the possibility of "therapeutic help" through the adaptive social communication that is made possible by this process.

We will begin with the idea of personal narratives. By this we mean the ways in which we understand ourselves in relation to the world, our history and our relationships. Each of us will have different and possibly competing narratives by which we understand who we are and what is going on for us at any one time. From a phenomenological perspective, these represent various ways of *Dasein* ("Being-in-the-world") [15, 16]. We might be more immediately aware of and preoccupied by some of these narratives than others, but they are all experientially accessible characterizations of one's self, and together they constitute one's "sense of self". That all individuals have a personal narrative, an imagined sense of self evidenced by our experiences, has long been recognized by phenomenologists [17], and the biological reality of this has been impressively demonstrated by the research summarized by Northoff and Huang [18]. We suggest that one of the key experiences that makes therapeutic change possible is the recognition of these personal narratives – notably, the minor, more complex narratives as well as the predominant story of one's self that may be present. The therapist's recognition and articulation of these narrative threads is a significant part of the therapeutic process because, we have argued, this experience is a potent ostensive cue for the stimulation of epistemic trust (defined as openness to the reception of social knowledge that is regarded as personally relevant and of generalizable significance) [3, 19]. The recognition of agency suggested by the explicit understanding and elucidation of another's personal narrative signals a shared intentionality. In brief, if an individual experiences himself or herself as being understood, he/she will be inclined to learn from the person who has shown that they understand him/her. This will include learning about oneself but also about others and about the environment in which one lives – mostly significantly, how to navigate the social and cultural environment with all its complexities and challenges. In a highly cited paper, James Strachey introduced the concept of "mutative interpretation" – the therapist acting as an auxiliary superego helping the patient to recognise impulses or elements in him or herself to produce change in the patient's mental organization [20]. We suggest that the conscious and explicit articulation of difficult, not consciously recognised narratives by the therapist acting as "superego" is such a powerful tool, in terms of the evolutionary thinking we have described here, because it enables the patient to develop their capacity for social learning. The concept of the auxiliary superego might be understood as holding an epistemic authority to which we are highly primed, in evolutionary terms, to respond.

In order to safely depend on others to learn about reality, we need to be able to identify those who are reliable sources of information. The young human needs to be able to distinguish trustworthy, benevolent and reliable sources of knowledge from those communicators who are either poorly informed or badly intentioned. In either case, the latter are the purveyors of useless or deceptive information. Thus in order to ensure effective cultural knowledge-transfer via teaching, humans needed to evolve a reliable way of distinguishing trustworthy sources of knowledge. Trust in knowledge (which we call epistemic trust following Sperber, [21, 22]) is at the heart of what it means to be a human. All young humans are at the mercy of a knowledge differential, uncertain about the trustworthiness of the information they are about to receive, but are able to rapidly establish epistemic trust in order to benefit from a rapid and efficient system of knowledge transfer. Epistemic vigilance is the self-protective suspicion towards potentially damaging, deceptive, or inaccurate information [21]. The capability for vigilance as well as a mechanism for selectively circumventing it must be profound and deeply etched into our human origins.

It is clear that the absence of epistemic trust would deeply disadvantage an individual in most social contexts. The loss of this key process for the efficient acquisition of cultural knowledge has significant implications for social functioning. The individual may become limited in their ability to update their understanding of potentially rapidly changing social situations and would appear inflexible or even rigid in the face of social change. Why would an individual fail to experience epistemic trust even in situations where trust was warranted—that is, where their personal narrative was appreciated? There are two obvious reasons. First, adversity and deprivation, when tantamount to trauma, can generate chronic mistrust by inhibiting imagination, creating an overarching avoidance of mentalizing and an almost phobic avoidance of mental states, leaving the individual deeply vulnerable in most social situations. We use imagination here to refer to the capacity to form a second order of representation, tying back to the original Latin definition of imaginary ("to form an image, represent"). Even in the absence of such a pervasive failure of imagination, inadequate mentalizing may lead the traumatized individual to be biased in their perception of social reality [23-26] and misrepresent how others represent them, leading them to feel persistently misunderstood. Secondly, the long-term outcome of epistemic mistrust secondary to the failure of imagination we will describe here may create problems for individuals who have distorted personal narratives that generate inaccurate views of the self, so that even an accurate perception of one's personal narrative by others is not

experienced as a match, and a painful experience of interpersonal alienation persists. Conversely, in yet other instances, deprivation and trauma may generate inappropriate trust. We understand such excessive epistemic credulity as triggered by a hyperactive or unmoored social imagination generating a personal narrative that is so diffuse that the individual concerned is unable to judge whether another person's perception of them is accurate. Excessive credulity results as all personal narratives feel as if they "fit" sufficiently for trust to be generated, making the person vulnerable to exploitation. Of course, limited imagination may cause profound misperceptions of the other's representations of one's personal narrative, and an illusory fit is created where none in reality exists. There may be many other possibilities.

We suggest that in all these permutations an individual's social experience leads them to encounter problems in learning from others, which in turn creates significant problems in adaptation when they attempt to adjust to a frequently challenging and changing social world. But all these permutations possess a shared quality that derives from the individual's difficulty in being able to work with other minds to rectify their perception of their own mind in relation to the social environment in a way that delivers affect regulation, and helps to shore up executive function. Individuals in a state of heightened epistemic mistrust will not benefit from the access to other people's minds that could serve to regulate their own imaginative activity. Without the social metric that epistemic trust enables, the imagination may "run riot," and go substantially beyond the shared reality that people ultimately must agree on in order to collaborate. Difficulties with reaching agreement with other minds are characteristic of many forms of mental health disorder, and personality disorder perhaps most paradigmatically. The system of cultural transmission that humans have evolved requires imagination (of which mentalizing is one aspect) in order to establish trust; however, the transmission of knowledge that follows places a constraint on the imagination to ensure that there is an agreed version of reality. Being able to mentalize one another makes it possible to have a collectively agreed imagination which makes human cooperation possible [27]. The significance of epistemic trust in relation to our model of psychopathology is therefore that it enables the individual to align their social imagination with the prevailing social reality in an adaptive way, creating the foundation for the intergenerational transmission of ideas and the creation of social networks that in turn support culture.

We have recently suggested that effective psychotherapeutic practice taps into this human capacity for imagination, and that psychopathology and disruptions in mentalizing involve dysfunctional imaginative processes that obstruct the individual's salutogenic exposure to social communication [28, 29] [for a discussion of the idea of salutogenesis as an approach that considers the factors supporting health see 30]. In particular, the interpersonal component of this process is essential. In recognising and jointly considering the subjective experience of the individual, it becomes endowed with a conscious significance. This recognition by consciousness is valuable because it creates the conditions for epistemic trust and the possibility of adaptive social communication and learning with others.

The three aspects of the communication process within psychotherapy

We have described the processes that underpin effective psychotherapy elsewhere (we have previously labelled these the three communication systems [29]). Here we would like to approach these processes in terms of phenomenological experience, i.e. in terms of the subjective experience underlying them. In phenomenological terms, it is an account of change in relation to experiences involving the *Eigenwelt* ("the own world"), the *Mitwelt* ("the with world", involving interpersonal relatedness), and the broader *Umwelt* ("the around world'"), to use the famous phenomenological psychiatrist Binswanger's terminology [15]. Through this account of the communicative unfolding of psychotherapy, we hope to elucidate how it is embedded in metacognitive processes about the self in relation to perceptual social reality. In addition to this, we would like to bring in social reality in a more immediate sense, in considering the impact of the phenomena of socio-economic deprivation, inequality and social isolation on psychopathology.

1) The epistemic match

All evidence-based psychotherapies provide a coherent framework that enables the patient to examine the issues that are deemed to be central to him/her, according to a particular theoretical approach, in a safe and low-arousal context. Psychotherapeutic models differ in detail, but generally work – directly or indirectly – to develop strategies to handle how one thinks and feels with regard to oneself (the *Eigenwelt*) and restructure thinking about interpersonal relationships (the *Mitwelt*). Perhaps more importantly, however, all evidence-based psychotherapies

provide for the patient a model of mind and an understanding of their disorder, as well as a hypothetical appreciation of the process of change, that are accurate enough for the patient to feel recognized and understood as an agent. Any therapeutic model – i.e., understanding of the causes of the problem and their possible resolution – can only be effective insofar as it results in the feeling of being mirrored in a marked way, which leads to the feeling of being understood. This, in our view, is one of the most powerful human experiences leading to the restoration of feelings of agency and selfhood. These experiences in turn lead to recovery of mentalizing and epistemic trust.

In essence, we suggest that such explanations and suggestions may be seen as ostensive cues that signal to the patient the relevance to them of the information that is being conveyed. Csibra and Gergely take the concept of "ostensive cues" [31] – discussed originally by Bertrand Russell [32], but extensively used by Sperber and Wilson [33] – to mean that certain signals are employed by an agent and prepare the addressee for the intent of the agent to communicate. They are signals designed to trigger epistemic trust. Examples of ostensive cues are eye contact, eyebrow raising, contingent reactivity and infant-direct speech (motherese). The particular process of ostensive cueing in psychotherapy – via the therapist's rich and careful mentalizing of the patient – is important because it allows the patient to reduce his/her epistemic hypervigilance as he/she increasingly sees the model's relevance to his/her own state of mind. Thus, acquiring new skills and learning new and useful information about oneself, as well as doubtless being useful in its own right, has the nonspecific effect of creating epistemic openness. This openness makes it easier for the patient to learn the specific suggestions conveyed within the model. A virtuous cycle is created: the patient "feels" the personal truth of the content conveyed within the therapeutic model, which, because it is accurate and helpful, generates epistemic openness. The growth of epistemic trust allows the patient to take in further information that also serves to reassure and validate him/her.

As will be explained in more detail below, in our discussion of the third aspect of communication, we need to take into account the role of wider social system (the "Umwelt") in generating a feeling of subjective alienation and epistemic trust. The first therapeutic task at hand is therefore to recognise this experience to restore a feeling of subjectivity and epistemic trust. Any theoretical model – no matter how robust or accurate it may be – will be completely powerless in patients with feelings of subjective alienation and epistemic mistrust unless this task is

achieved. For the sake of narrative clarity, we are conveying this as a linear progression. The clinical reality is that it is a task that will in most cases need to be revisited and will overlap with the processes unfolding as part of communication aspect 2 and 3. As noted, this alienation might originate in particular psychological problems, for example a severely depressed patient feeling completely hopeless and beyond help, but an individual's social circumstances might indeed be highly alienating (acute social deprivation is a case in point) – for such individuals it might be adaptive to trust those who claim to offer help. The approach to psychopathology that we have described here is an evolutionary one, which regards many forms of disorder as originating as a form of adaptation to social circumstances [28, 29]. The "social alienation" associated with inhabiting a more broadly non-mentalizing social system might be understood as a generalized breakdown in epistemic trust. A recognition of the presence of these wider processes may be a necessary extension to the dyadic emphasis of the therapeutic approach in order to give what the therapist is communicating a sense of phenomenological reality. It is through the therapist's understanding of, adaptation to and effective marked mirroring of the patient's vigilant stance and its origins that the work of the first aspect of communication is achieved.

2) Improving mentalizing

As noted above, through passing on knowledge and skills that feel appropriate and helpful to the patient, the clinician is actively recognizing the patient's agency. The clinician's presentation of information that is personally relevant to the patient serves as a form of ostensive cueing that conveys the impression that the clinician seeks to understand the patient's perspective; this in turn enables the patient to listen to and hear the clinician's intended meaning. In effect, the clinician is demonstrating how he/she engages in mentalizing in relation to the patient. It is important that in this process both patient and clinician come to see each other more clearly as intentional agents. For example, when the clinician shows that his/her mind has been changed by the patient, her/she gives agency to the patient and increases his/her faith in the value of social understanding. The context of an open and trustworthy social situation facilitates the achievement of a better understanding of the beliefs, wishes and desires underpinning the actions of others and of the self. This allows a more trusting relationship to develop between clinician and patient. Ideally, the patient's feeling of having been sensitively responded to by the clinician opens a second virtuous

cycle in interpersonal communication in which *the patient's own capacity to mentalize is regenerated.* This we believe constitutes an important turning-point in all types of psychotherapy – when the patient begins to develop genuine interest and curiosity in his own mind, that of others around him including the therapist.

Understanding the patient's subjectivity is vital to this process, as the patient's self-discovery as an active agent occurs through the social interchange where they experience themselves as an agent in the way their clinician thinks of them – it could be said that they 'find themselves in the mind of the clinician'. It is also vital to a further function of therapy: the rekindling of the patient's wish to learn about the world, including the social world. We believe that this is a complex and non-linear process, but it can be summarized briefly as follows: the insight obtained in therapy, whatever its content, creates or recreates the potential for the patient to have a learning experience, which in turn makes other similar learning experiences more productive because it *enables the patient to adopt a stance of learning from experience by increasing their capacity to mentalize*.

The benefit of improved mentalizing as part of the social process of psychological therapy feeds back to increasing epistemic openness in two ways. Firstly, with improved mentalizing the individual becomes more sensitive and accurate in identifying their personal narrative (their phenomenological experience) in the implicit presentation of them by the therapist. Secondly, improved mentalizing also generates enhanced and more nuanced self experience that in turn facilitates the process of self recognition in the social context of therapy. In both these ways, increasingly robust mentalizing will serve to gradually improve communication between therapist and patient and enable him/her to benefit from new knowledge which the therapeutic process brings.

As an example, phenomenological psychiatry has linked depression to a disturbance in the experience of time [34]: past, present, and future do not have the same differentiated meaning for the depressed patient as it has for individuals without disturbed mood, but all feel equally painful and immovable [see also 35]. This experience of being locked in the "specious present" [36] leads to feelings of helplessness and hopelessness, and disturbances in the experience of both time. Yet, what we typically observe in this phase of treatment is that when the depressed individual recognises, in dialogue with a reflective therapist, that this feeling is borne out of psychic equivalence (i.e. the conviction that what one thinks or feels is true), rather than being a true reflection of reality, it opens up the

possibility of recognising other ways of seeing the self in time. And this experience helps to open up the mind of the patient more generally to other, alternative ways of thinking and feeling about the self, other ways of *dasein* ("being-in-the-world').

3) The re-emergence of social learning outside therapy

The improved mentalizing that results from effective treatment brings about improved social relations and experiences outside the consulting room. Improved levels of trust and the breaking down of rigid ways of interpreting and responding to social experiences pave the way for the patient client to accumulate experiences of social interaction that are benign, or that are at least manageable in terms of maintaining resilient mentalizing. This creates another virtuous circle in which more balanced and robust mentalizing generate and support deeper, wider, and increasingly meaningful access to social information and social networks. This final, critical stage of social learning beyond therapy is of course contingent dependent on the individual's social environment being benign, or at least "benign enough"'. Therapeutic change can only be sustained, according to this thinking, if the patient client is able to use, and even to change (through the seeking out of more mentalizing relationships), their social environment in a way that allows them to continue to relax epistemic hypervigilance and foster their mentalizing strengths.

We suggest that for individuals who are enduring mental health difficulties in the context of greater socio-economic inequality and deprivation, aspect 3 of communication process may be of heightened significance. We would like to explain the relevance and significance of this process here, as it speaks to the phenomenological experience of therapeutic change, and the importance of considering the *Umwelt*, the broader sociocultural context in conceptualizing therapeutic change.

We know that at least some risk factors cut a swathe through the complexity of individual diagnoses, and perhaps the most powerful of these is socioeconomic status [37, 38]. There is unequivocal evidence that social and economic inequality is strongly connected with mental ill health [39-41], and poverty is one of the best documented risk factors for both internalising and externalising problems [42]. It has been argued that the dominance of biological and individual psychological perspectives may have distracted from considering broader social

perspectives [43, 44]. And indeed, we are increasingly of the view that the role of these wider social systemic experiences – while abundantly clear to many clinicians working on the front line of mental health care – is theoretically under-accounted for in existing conceptualisations psychopathology. In our description of the aspects of communication in effective psychotherapy, we would like to propose a model that takes into account these wider social phenomena and their relationship to individual psychic distress.

We have evidence that individuals who are less socio-economically privileged tend to behave in more community and socially oriented ways in interpersonal trust experiments than more affluent individuals [45]. Less affluent individuals are more engaged with and dependent on their community; wealthier and more social protected individuals have a stronger perception of their self agency and as a result need to be less community focussed [46-48]. As a result, individuals functioning in a lower SES environment are also therefore more sensitive to their social environment, its reliability and how benign or supportive it may be [49, 50]. The flipside of this is that when the social environment is hostile or unsupportive, the individual may be more responsive to the meaning and significance of that, increasing the breakdown in social learning – resulting in what is recognised in sociological terms as social alienation. It is this effect, we posit, that contributes to the relationship between socio-economic forces and poor mental health outcomes. According to this thinking, the reason why inequality rather than absolute income level is so pernicious for mental health can be explained in terms of the sense of social vulnerability and breakdown associated with it [51, 52]. In summary, socioeconomic disadvantage is likely to be a powerful cause of mental disorder across diagnostic categories but its impact is moderated by interpretation of the meaning of disadvantage.

A similar complex pattern of how overarching social factors impact on individual risk emerges from studies of the association of ethnicity with mental disorder. Amongst children, subjective wellbeing is either unrelated to ethnicity [e.g., 53] or its association is modest [e.g., 54]. Racial discrimination is however a powerful predictor of general psychopathology [55] with the most powerful associations observed for depression and anxiety [56, 57] and conduct problems [58, 59]. As one might expect, from the broad range of mental disorders found at greater prevalence with groups subject to racial prejudice, it is once again the shared transdiagnostic component of psychopathology which correlates with the experience of racial discrimination [60].

The implication of this thinking is that this third aspect of communication is of particular significance to patients who are relatively socially powerless or deprived. The challenge for the therapist is to support the patient in building or sustaining mentalizing social relationships in what might be more challenging environments. But it is only if such conditions can be created that the hope of effective therapeutic changed can be realistically entertained. This is because it is only once the patient encounters the phenomenological reality of such an environment that – for entirely adaptive reasons – we can reasonably expect them to become open to the social learning and sustenance made possible by increased mentalizing capacity.

Conclusion

It is the recovery of the capacity for social information exchange that, we feel, may be at the heart of effective psychotherapies. As clinicians we often assume that what happens in the consulting room is the primary driver of change, but experience shows us that change is also brought about by what happens beyond therapy, in the person's social environment. Studies in which change was monitored session by session have suggested that the patient–clinician alliance in a given session predicts change in the next [61]. This indicates that the change that occurs between sessions is a consequence of changed attitudes to learning engendered by therapy, influencing the patient's behaviour between sessions.

The factors associated with 'therapies that work' create experiences of truth – subjectively felt truth – which in turn encourage the patient to learn more. In this process, via a non-specific pathway, the patient's capacity to mentalize is fostered. Both of these systems would be expected to lead to symptomatic improvement. Improved mentalizing and reduced symptomatology both improve the patient's experiences of social relationships. However, it is likely that these new and improved social experiences, rather than just what happens within therapy, serve to erode the epistemic hypervigilance that has previously prevented benign social interactions from changing the patient's experience of themselves and of the social world. *Meaningful change is thus possible only if the person can use their social environment in a positive way (and if the social environment is sufficiently supportive to allow this to happen)*. For this to happen, recognition of self-agency is key, and this recognition is best achieved through the ostensive cues that are provided by feeling appropriately mentalized by another person. For the social environment

to be accurately interpreted so that it can provide opportunities for new learning, mental state understanding of others' actions and reactions is critical – and only improved mentalizing will achieve this. Hence, as in the phenomenological tradition, we believe that changes in the subjective experience of the self, particularly those that take place in relation to others that we imbue with epistemic trust, are at the heart of therapeutic change.

Acknowledgments

All authors contributed equally to the writing of this paper.

Disclosure Statement

Peter Fonagy and Patrick Luyten have been involved in the training and dissemination of mentalization-based treatments and hold research grants on mentalizing.

Funding Sources

Peter Fonagy is in receipt of a National Institute for Health Research (NIHR) Senior Investigator Award (NF-SI-0514-10157), and was in part supported by the NIHR Collaboration for Leadership in Applied Health Research and Care (CLAHRC) North Thames at Barts Health NHS Trust. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

References

1. Buhler, K.E., *Existential analysis and psychoanalysis: specific differences and personal relationship between Ludwig Binswanger and Sigmund Freud.* American Journal of Psychotherapy, 2004. **58**(1): p. 34-50.
2. Chalmers, D.J., *Facing up to the problem of consciousness* Journal of Consciousness Studies, 1995. **2**(3): p. 200-219.
3. Fonagy, P., P. Luyten, and E. Allison, *Epistemic petrification and the restoration of epistemic trust: A new conceptualization of borderline personality disorder and its psychosocial treatment.* Journal of Personality Disorders, 2015. **29**(5): p. 575-609.
4. Safran, J.D., J.C. Muran, and C. Eubanks-Carter, *Repairing alliance ruptures*, in *Psychotherapy relationships that work: Evidence-based responsiveness*, J.C. Norcross, Editor. 2011, Oxford University Press: New York, NY. p. 224-238.
5. Mercier, H. and D. Sperber, *The enigma of reason: A new theory of human understanding*. 2017, London: Allen Lane.
6. Mahr, J. and G. Csibra, *Why do we remember? The communicative function of episodic memory.* Behavioral and Brain Sciences, 2017: p. 1-93.
7. Fonagy, P. and E. Allison, *Psychic reality and the nature of consciousness.* International Journal of Psychoanalysis, 2016. **97**(1): p. 5-24.

8. Dennett, D.C., *Consciousness Explained*. 1991, Boston: Little Brown.
9. Auerbach, J.S. and S.J. Blatt, *Self-representation in severe psychopathology: The role of reflexive self-awareness.* Psychoanalytic psychology, 1996. **13**: p. 297-341.
10. Klein, R.G., *Paleoanthropology. Whither the Neanderthals?* Science, 2003. **299**(5612): p. 1525-7.
11. Klein, R.G., *Ice-age hunters of the Ukraine* 1973, Chicago, IL: University of Chicago Press.
12. Debbané, M., et al., *Attachment, neurobiology, and mentalizing along the psychosis continuum.* Frontiers in Human Neuroscience, 2016. **10**: p. 406.
13. Debbané, M., et al., *Mentalization-based treatment in clinical high-risk for psychosis: A rationale and clinical illustration.* Journal of Contemporary Psychotherapy, 2016. **46**(4): p. 217-225.
14. Fotopoulou, A. and M. Tsakiris, *Mentalizing homeostasis: The social origins of interoceptive inference.* Neuropsychoanalysis, 2017. **19**(1): p. 3-28.
15. Needleman, J., ed. *Being-in-the-world: Selected papers of Ludwig Binswanger* 1963, Basic Books: New York.
16. Heidegger, M., *Being and Time*. 2010, New York: State University of New York Press.
17. Sass, L., et al., *EAWE: Examination of Anomalous World Experience.* Psychopathology, 2017. **50**(1): p. 10-54.
18. Northoff, G. and Z. Huang, *How do the brain's time and space mediate consciousness and its different dimensions? Temporo-spatial theory of consciousness (TTC).* Neuroscience & Biobehavioral Reviews, 2017. **80**: p. 630-645.
19. Fonagy, P. and E. Allison, *The role of mentalizing and epistemic trust in the therapeutic relationship.* Psychotherapy, 2014. **51**(3): p. 372-380.
20. Strachey, J., *The nature of the therapeutic action of psychoanalysis.* International Journal of Psycho-Analysis, 1934. **15**: p. 275-292.
21. Sperber, D., et al., *Epistemic vigilance.* Mind & Language, 2010. **25**(4): p. 359-393.
22. Wilson, D. and D. Sperber, *Meaning and relevance*. 2012, Cambridge, UK: Cambridge University Press.
23. Germine, L., et al., *Childhood adversity is associated with adult theory of mind and social affiliation, but not face processing.* PLoS One, 2015. **10**(6): p. e0129612.
24. Kay, C.L. and J.M. Green, *Social cognitive deficits and biases in maltreated adolescents in U.K. out-of-home care: Relation to disinhibited attachment disorder and psychopathology.* Development and psychopathology, 2016. **28**(1): p. 73-83.
25. Armelius, B.A. and T.H. Andreassen, *Cognitive-behavioral treatment for antisocial behavior in youth in residential treatment.* Cochrane Database of Systematic Reviews, 2007(4): p. CD005650.
26. Cicchetti, D. and W.J. Curtis, *An event-related potential study of the processing of affective facial expressions in young children who experienced maltreatment during the first year of life.* Development and Psychopathology, 2005. **17**(3): p. 641-77.
27. Tomasello, M., *Great apes and human development: A personal history.* Child Dev Perspect, 2018.
28. Fonagy, P., et al., *What we have changed our minds about: Part 1. Borderline personality disorder as a limitation of resilience.* Borderline Personality Disorder and Emotion Dysregulation, 2017. **4**: p. 11.
29. Fonagy, P., et al., *What we have changed our minds about: Part 2. Borderline personality disorder, epistemic trust and the developmental significance of social communication.* Borderline Personality Disorder and Emotion Dysregulation, 2017. **4**: p. 9.
30. Antonovsky, A., *Unraveling the mystery of health: How people manage stress and stay well*. 1987, San Francisco, CA: Jossey-Bass.
31. Csibra, G. and G. Gergely, *Natural pedagogy.* Trends in Cognitive Sciences, 2009. **13**(4): p. 148-53.
32. Russell, B., *An inquiry into meaning and truth*. 1940, London, UK: Allen & Unwin.
33. Sperber, D. and D. Wilson, *Relevance: Communication and cognition*. 2nd ed. 1995, Malden, MA: Blackwell.
34. Ey, H., *Etudes psychiatriques. Tome III. Structure des psychoses aigues et déstructuration de la conscience*. 1954, Paris: Desclée de Brouwer.
35. Kernberg, O.F., *The destruction of time in pathological narcissism.* International Journal of Psychoanalysis, 2008. **89**(2): p. pp.
36. James, W., *Principles of psychology*. 1890, New York, NY: Henry Holt.
37. Carod-Artal, F.J., *Social determinants of mental health*, in *Global mental health: Prevention and promotion*, S. Bährer-Kohler and F. Carod-Artal, Editors. 2017, Springer International: Cham, Switzerland.

38. Allen, J., et al., *Social determinants of mental health.* International Review of Psychiatry, 2014. **26**(4): p. 392-407.
39. Pickett, K.E. and R.G. Wilkinson, *Inequality: An underacknowledged source of mental illness and distress.* British Journal of Psychiatry, 2010. **197**(6): p. 426-8.
40. Platt, S., S. Stace, and J. Morrissey, *Dying from inequality: Socioeconomic distadvantage and suicidal behaviour*. 2017, London, UK: Samaritans.
41. Silva, M., A. Loureiro, and G. Cardoso, *Social determinants of mental health: A review of the evidence.* European Journal of Psychiatry, 2016. **30**(4): p. 259-292.
42. Yoshikawa, H., J.L. Aber, and W.R. Beardslee, *The effects of poverty on the mental, emotional, and behavioral health of children and youth: Implications for prevention.* American Psychologist, 2012. **67**(4): p. 272-84.
43. Smith, M.A., *A fine balance: Individualism, society and the prevention of mental illness in the United States, 1945–1968.* Palgrave Communications, 2016. **2**(1).
44. Shim, R., et al., *The social determinants of mental health: An overview and call to action.* Psychiatric Annals, 2014. **44**(1): p. 22-26.
45. Dubois, D., D.D. Rucker, and A.D. Galinsky, *Social class, power, and selfishness: When and why upper and lower class individuals behave unethically.* Journal of Personality and Social Psychology, 2015. **.108**(3): p. pp.
46. Keltner, D., D.H. Gruenfeld, and C. Anderson, *Power, approach, and inhibition.* Psychological Review, 2003. **110**(2): p. 265-84.
47. Lee, F. and L.Z. Tiedens, *Is it lonely at the top?: The independence and interdependence of power holders.* Research in Organizational Behavior, 2001. **23**: p. 43-91.
48. Overbeck, J.R. and V. Droutman, *One for all: Social power increases self-anchoring of traits, attitudes, and emotions.* Psychological Science, 2013. **24**(8): p. 1466-76.
49. Gruenfeld, D.H., et al., *Power and the objectification of social targets.* Journal of Personality and Social Psychology, 2008. **95**(1): p. 111-27.
50. van Kleef, G.A., et al., *Power, distress, and compassion: Turning a blind eye to the suffering of others.* Psychological Science, 2008. **19**(12): p. 1315-22.
51. Wilkinson, R.G. and K.E. Pickett, *The enemy between us: The psychological and social costs of inequality.* European Journal of Social Psychology, 2017. **47**(1): p. 11-24.
52. Pickett, K.E. and R.G. Wilkinson, *Income inequality and health: A causal review.* Social Science & Medicine, 2015. **128**.
53. Huebner, E.S., et al., *A review of the brief multidimensional students' life satisfaction scale.* Social Indicators Research, 2006. **79**(3): p. 477-484.
54. Bradshaw, J., et al., *Children's subjective well-being: International comparative perspectives.* Children and Youth Services Review, 2011. **33**(4): p. 548-556.
55. Clark, R., et al., *Racism as a stressor for African Americans. A biopsychosocial model.* American Psychologist, 1999. **54**(10): p. 805-16.
56. Banks, K.H., L.P. Kohn-Wood, and M. Spencer, *An examination of the African American experience of everyday discrimination and symptoms of psychological distress.* Community Mental Health Journal, 2006. **42**(6): p. 555-70.
57. English, D., et al., *Neighborhood racial composition, racial discrimination, and depressive symptoms in African Americans.* American Journal of Community Psychology, 2014. **54**(3-4): p. 219-28.
58. Brody, G.H., et al., *Perceived discrimination, serotonin transporter linked polymorphic region status, and the development of conduct problems.* Development and Psychopathology, 2011. **23**(2): p. 617-27.
59. Brody, G.H., et al., *Perceived discrimination and the adjustment of African American youths: A five-year longitudinal analysis with contextual moderation effects.* Child Development, 2006. **77**(5): p. 1170-89.
60. Liu, J., et al., *Risk and protective factors for comorbid internalizing and externalizing problems among economically disadvantaged African American youth.* Development and Psychopathology, 2017. **29**(3): p. 1043-1056.
61. Falkenstrom, F., F. Granstrom, and R. Holmqvist, *Therapeutic alliance predicts symptomatic improvement session by session.* Journal of Counseling Psychology, 2013. **60**(3): p. 317-28.