ORIGINAL PAPER

Mentalization and Psychosis: A Rationale for the Use of Mentalization Theory to Understand and Treat Non-affective Psychotic Disorder

J. G. Weijers[1,2,7,8] · C. ten Kate[1] · M. Debbané[3] · A. W. Bateman[4] · S. de Jong[5] · J. -P. C. J. Selten[1,2] ·
E. H. M. Eurelings-Bontekoe[6]

Abstract

Social functioning can be severely impaired in non-affective psychotic disorder (NAPD). Current models of psychosis pathogenesis do not tend to focus on social dysfunction and pharmacological treatment fails to ameliorate it. In this article, we propose that mentalization theory provides a valuable contribution to the understanding and treatment of NAPD. Impaired mentalizing may contribute to both positive and negatives symptoms as well as social dysfunction observed in NAPD. Furthermore, impaired mentalizing may help explain the relation between childhood abuse, insecure attachment and psychosis. Mentalization based treatment may contribute to the functional recovery of NAPD patients as it targets the social cognitive processes underlying social interaction. The article includes a description of the principles of MBT in general, specific characteristics of using MBT with patients with NAPD and a clinical vignette to illustrate these principles.

Keywords Psychosis · Schizophrenia · Attachment theory · Mentalizing · Epistemic trust · Mentalization based treatment

Part I: A Mentalization Based Model of Psychosis Pathogenesis

Introduction

A recent, dominant model of psychosis pathogenesis (Howes and Murray 2014) suggests that many factors contribute to the etiology of psychosis, such as variant genes, stress, neuroinflammation, dysregulated activity in the hypothalamic–pituitary–adrenal axis, and developmental insults (e.g., pre- and perinatal complications). It is furthermore held that the assignment of aberrant salience (i.e., motivational value) to mundane stimuli due to a sensitized mesolimbic dopamine system is the final common pathway through which these variables increase psychosis risk. Additionally, there is increased consensus that misattribution processes in the form of cognitive biases are also involved in the etiology of psychosis. For example, hallucinations do not solely involve aberrant experience but also involve the misattribution of such experiences to outside sources

✉ J. G. Weijers
j.weijers@rivierduinen.nl

C. ten Kate
c.tenkate@rivierduinen.nl

M. Debbané
martin.debbane@unige.ch

A. W. Bateman
anthony.bateman@ucl.ac.uk

S. de Jong
steven.dejong@lentis.nl

J. -P. C. J. Selten
J.Selten@rivierduinen.nl

E. H. M. Eurelings-Bontekoe
EURELING@fsw.leidenuniv.nl

[1] Rivierduinen Institute for Mental Health Care, Leiden, The Netherlands

[2] Maastricht University, Leiden, The Netherlands

[3] Université de Genève, Geneva, Switzerland

[4] University College London, London, United Kingdom

[5] Lentis Psychiatric Institute, Groningen, The Netherlands

[6] Leiden University, Leiden, The Netherlands

[7] Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Maastricht University, Maastricht, The Netherlands

[8] GGZ Rivierduinen, Sandifortdreef 19, 2333 ZZ Leiden, The Netherlands

(Aleman and Larøi 2008). Thirdly, evidence suggests that socio-developmental adversity plays a vital role in the development of psychosis through the sensitization of the mesolimbic dopamine system (e.g., Selten et al. 2013).

As such, the recent conceptualization of psychosis (Howes and Murray 2014) offers a much needed integration of different neurodevelopmental, biological, and cognitive models to account for psychotic or "positive" symptoms, such as delusions or hallucinations. However, non-affective psychotic disorders (NAPD[1]) like schizophrenia, brief psychotic disorder and schizoaffective disorder, also involve negative symptoms such as lack of initiative or flattened affective expression and social dysfunction (Green et al. 2019), which are less prominently featured in most current etiological models of psychosis. Because mentalization theory focuses on the development of social cognitive processes, it may offer a valuable contribution to current models of NAPD.

Social Dysfunction and Social Cognition in NAPD

Research shows that premorbid social dysfunction and adversity predict poor clinical and functional outcome in NAPD (e.g. White et al. 2009; Velthorst et al. 2010; Stumbo et al. 2015). Whereas positive symptoms often can be treated adequately with antipsychotic medication, social dysfunction cannot (Pinkham et al. 2003), often creating a wide gap between the severity of symptoms and dysfunction in patients with chronic NAPD (e.g., Birchwood et al. 2013). However, a recent review showed that functional recovery *is* possible, with the rate for recovery of adequate social functioning lying between 14 and 53% (Leonhardt et al. 2017). Rather than pharmacological factors however, *social* factors such as feeling connected (Eisenstadt et al. 2012; Hendryx et al. 2009) and experiencing social support (Norman et al. 2013; Thomas et al. 2016) seem instrumental to subjective and objective functional recovery. As a result, the development of nonpharmacological, adjunctive treatments that improve functioning in NAPD has been an important goal in recent psychosis research (e.g., Holthausen et al. 2007).

Social cognition is the multifaceted domain of cognitive and affective processes that underlie social interaction. Given that social cognitive deficits are widely linked to poor social functioning in patients with NAPD (see Fett et al. 2011), they have garnered increased interest as a treatment target. Many lines of research opt to focus on specific aspects of social cognition such as facial affect recognition or "theory of mind," i.e., the ability to infer mental states of

others (Green et al. 2019). However, it can be argued that well-developed social cognition comprises the successful synthesis of its different aspects. Increasing the successful recognition of facial affective expression, may help to recognize *that* someone is angry, but does not necessarily mean that one is better able to interpret *why* a person is angry. Recent lines of research have therefore taken a more 'holistic' approach, proposing that underlying social cognitive deficits is an impaired ability to think about thoughts and emotions (Semerari et al. 2003) also termed mentalization (Fonagy 1989) or metacognition (Lysaker et al. 2005). While this article mainly focuses on mentalizing, we will also touch upon metacognition, given the conceptual overlap of these constructs.

Social Cognition, Mentalization and Metacognition

To understand the constructs of mentalizing and metacognition, we first need to look at theoretical accounts of how people come to understand minds. Until recently, two theories polarized the field of social cognition (Keysers and Gazzola 2007). Simulation theory holds that people develop an experiential understanding of others through the simulation of others' experiences, through mirror neurons and shared sensory-affective circuits (Gallese 2005). The intended meaning of actions is viewed to be understood implicitly, through shared feeling. 'Theory theory' on the other hand, holds that we come to understand others through the cognitive inference of "rules" from observed social situations, thus attributing theoretical beliefs and intentions to them (e.g. Baron-Cohen et al. 1986). These rules are thought to be updated if the models do not precisely predict current situations.

It can be argued that both metacognition and mentalization, bridge the gap between these two positions. Mentalizing—defined as a form of "imaginative mental activity through which behavior is interpreted in terms of mental states like needs, feelings, beliefs and goals" (Bateman and Fonagy 2004)—can be understood as the dynamic interplay between implicit, reflexive, affect-oriented processes and explicit, reflective, cognition-oriented processes. Metacognition refers to a "spectrum of mental activities by which persons form integrated ideas about their own minds and those of others" (Lysaker and Dimaggio 2014). Metacognition comprises discrete activities such as affect recognition and more synthetic activities such as the integration of experiences into overarching narratives. Both concepts suggest that we do not just experience another's mental state through shared emotions—this would not differentiate us much from macaques (Gallese 2005)—but that we use higher-order cognitive processes to transform discrete experiences into conscious representations that can be used to improve our understanding of ourselves as well as others. However, the

---

[1] This paper will focus exclusively on non-affective psychotic disorders (NAPD), because affective psychotic disorders may have a different pathogenesis.

concept of mentalization is anchored in psychoanalytic and particularly attachment theory, specifying particular preconditions for the development of mentalizing ability, namely, a parent's capacity to reflect on the mind of the child. In contrast, metacognition is posited as an integrative model that is not specifically linked to a developmental theory.[2]

Mentalizing Impairments and NAPD

A growing body of research has observed a range of deficits in cognitive and affective aspects of mentalizing and metacognition in NAPD. Patients with NAPD tend to have an impaired ability to infer the mental states of others (for overviews see Harrington et al. 2005; Sprong et al. 2007) and understand others' emotional expressions (O'Driscoll et al. 2014). They also tend to have trouble recognizing their own internal sensory-affective experience (Brunelin et al. 2007) and show difficulty verbalizing such experience (Trémeau 2006). Frith (1992) was one of the first to suggest that an impaired capacity to experience and represent mental states of self and others is tied to specific clinical symptoms of NAPD. For example, delusions of persecution and reference are, almost by definition, failures to correctly represent the minds of others. Additionally, misunderstanding that certain gestures, expressions, and intonations convey mental states, may cause what clinicians perceive to be flattened affective expression (Frith 1992). An impaired ability to infer the mental states of others is also likely to severely complicate social interaction, as it leads to difficulty understanding its unwritten mores. This may make social interaction overwhelmingly complex and in turn may lead to withdrawal in patients or ostracism from peers. Indeed, mentalizing deficits seem related to both positive (e.g. Hasson-Ohayon et al. 2018; Mcleod et al. 2014) and negative symptoms (Mcleod et al. 2014; Weijers et al. 2018).

Impaired Embodied Mentalizing in NAPD

Debbané et al. (2016) further expanded upon Frith's hypothesis by suggesting that patients with NAPD suffer specifically from problems with 'embodied' mentalizing. Embodied mentalizing involves the ability to consciously detect and identify sensory-affective signals coming from one's body and to critically think about them. Indeed, NAPD patients tend to make errors in the detection and identification of self-generated events, or 'source monitoring errors' (see Brookwell et al. 2013 for an overview). Problems in detecting and accurately representing one's own sensory-affective

experience in turn cause problems in mentalizing (Lind and Bowler 2009; Kantrowitz et al. 2014). When embodied mentalizing becomes unbalanced, one may assign too much certainty to one's sensory-affective experiences as accurate representations of reality. If such experiences are not challenged and regulated by cognitive mechanisms such as reappraisal, they may come to be interpreted as equating with reality. When too much certainty is assigned to prior cognitive beliefs on the other hand, this runs the risk of misconstruing reality if such beliefs will fail to be updated by sensory-affective experience.

Epistemic Mistrust and Cognitive Biases in NAPD

Patients with NAPD not only have difficulty updating their beliefs on the basis of sensory-affective experience but often become unsusceptible to others' viewpoints. Psychosis often coincides with a loss of socially construed meaning ('common' sense in its literal meaning) and 'epistemic trust' (e.g. Pereira and Debbané 2018). Epistemic trust is defined as the "willingness to consider new knowledge from another person as trustworthy, generalizable, and relevant to the self" (Fonagy et al. 2015). According to the epistemic trust hypothesis, there are two ways in which people may come to accept new views as true: They can try to deduce by themselves whether such new knowledge offers an accurate model of reality, or they can rely on the knowledge of someone they deem both epistemically trustworthy and well-meaning (Fonagy et al. 2017).

A chronic lack of epistemic trust, sometimes referred to as epistemic hypervigilance, results in treating others as untrustworthy sources of information. It is thought to be pathogenic because it cuts individuals off from easy, interpersonal approaches to updating belief systems. Additionally, epistemic trust plays a central role in the sharing of cultural customs and values, because belonging to a certain cultural group gives rise to the expectation that one construes meaning of behavior in similar ways. Customs and values often have opaque functions but are nonetheless important to "fit in".

Contributors to Impaired Mentalizing and Psychosis

In the previous section, we argued that disruptions in mentalization and epistemic trust may be relevant to the pathogeny of NAPD. Here, we argue that childhood abuse may hinder the development of mentalizing and epistemic trust and thus fosters psychosis.

Many studies have shown that childhood abuse constitutes a significant risk factor to the emergence of psychosis (see Varese et al. 2012; Read et al. 2014 for overviews). Childhood abuse rates for those at ultra-high risk to develop psychosis may be as high as 86% (see Kraan et al. 2015

---

[2] However, it should be noted that attachment theory has received increased attention in recent metacognitive literature, see Aydin et al. (2016) for example.

for an overview), and prospective studies show that abused children are more likely to develop psychosis (e.g. Kelleher et al. 2013). Childhood abuse has also been widely shown to hinder the development of mentalizing ability (see Fonagy et al. 2017 for an overview) and impaired mentalizing has been shown to account for some of the relation between childhood abuse and negative symptoms in NAPD (Weijers et al. 2018).

Counter to nativist accounts of social cognition, mentalization theory proposes that mentalizing is initially formed in secure attachment relationships with caregivers (e.g., Kim 2015). Such relationships are characterized by a genuine interest in the child's mental state that is conveyed through ostensive cues (e.g., eye contact) and 'marked mirroring,' wherein emotions are imitated in a slightly exaggerated fashion, sometimes referred to as 'motherese' (Kim 2015). In this way, a child has his emotions "re-presented" back to him, which provides him with the opportunity to come into contact with other perspectives on his inner experiences (Fonagy and Allison 2014), thus enabling him to begin developing second-order representations of bodily feeling states. Although less emphasized, this hypothesis is also present in metacognitive theory, which stresses that the development of self-reflectivity is fostered by 'intersubjectivity', a communicative process between child and parent that revolves around the narration of inner experience (Lysaker et al. 2005). The experience of caregiver abuse on the other hand is thought to lead to a phobic avoidance of mentalizing in children (Kim 2015) "to protect themselves from acknowledging their caregiver's wishes to harm them" (Fonagy 1989). Inhibiting mentalizing in this manner may be the only means to negotiate the need for proximity and becoming overwhelmed by psychological pain.[3]

Childhood abuse may also contribute to poor mentalizing through the establishment of poor emotion regulation strategies. Attachment theory proposes that humans have an instinctual need to form close social bonds. In the face of danger or uncertainty, children display attachment behavior to promote safety and to regulate affect (Bowlby 1980). Especially, when primary caregivers are abusive, they are experienced as both a source of threat and a source of security to the child who is still dependent on them for nourishment and safety. In such cases, stress results in an alternation between strong impulses to escape and to seek care, increasingly escalating the levels of emotional dysregulation. Based on the quality of early attachment interactions a child develops expectations of relationships and how emotions will be co-regulated called 'attachment styles' (e.g., Berry et al. 2008). Perhaps unsurprisingly, attachment styles characterized by a distrust of others predominate in patients with psychosis. Between 70 and 89% of NAPD patients (dependent on the classification system used) are dismissing of attachment or attached in a disorganized fashion (e.g., Gumley et al. 2014). Estimates in the general population range between 20 and 25% for dismissing attachment (Ainsworth et al. 1978; Mickelson et al. 1997) and between 12% and 15% for disorganized attachment (e.g., Main & Solomon 1990). Moreover, greater insecure attachment in NAPD has been linked to greater difficulty regulating emotions (Owens et al. 2013).

Mentalizing in turn is vulnerable to excessive levels of emotional arousal (Nolte et al. 2013). Here, it is important to note that mentalizing is viewed as a complex 'higher-order cognitive process' in which functions of the prefrontal cortex, such as working memory and the direction of attention, are used to integrate pieces of information from different sensory, motor, and affective networks (Fonagy and Bateman 2016).[4] When overly stressed, the orchestration of the brain's activity by the prefrontal cortex is undermined and overtaken by subcortical structures, such as the amygdala, the nucleus accumbens, and the hypothalamus (Arnsten 2009) at the expense of higher order cognition.

Insecure Attachment and Epistemic Mistrust

Fonagy and Allison (2014) also maintain that epistemic trust, much like mentalizing capacity, is developed in secure attachment relationships. When a child finds himself accurately represented by a caregiver "as a thinking and feeling intentional being", this is thought to engender the secure feeling that the caregiver's intentions are benign, which helps the child more easily accept shared information as true. Indeed, the quality of the relationship between a child and a communicator determines the extent to which such information is accepted as truth (Lane and Harris 2015). However, children who repeatedly experience that their internal states are met with distorted or inaccurate caregiver responses, may develop a chronic mistrust regarding others' messages. Indeed, both insecure attachment styles and NAPD diagnoses are related to aspects of epistemic hypervigilance such

---

[3] By stipulating that childhood abuse and insecure attachment play a role in the pathogenesis of impaired mentalizing and psychosis, we in no way advocate a return to the days of the "schizophrenogenic" mother (see Hartwell 1996 for an overview). The contribution of peers to the development of mentalizing and secure attachment relationships seems substantial (see Choudhury et al. 2006; Sroufe 2005) and attachment styles are not as immutable as once thought (Mikulincer and Shaver 2012). Additionally, many forms of social adversity beyond child–caregiver attachment relationships predict risk of psychosis (Selten et al. 2013). Rather social relatedness throughout life seems crucial for mental health (Fonagy et al. 2017).

[4] Similarly, synthetic metacognition is thought to be the process combining "atoms of experience" into larger, overarching narratives (Hasson-Ohayon et al. 2018).
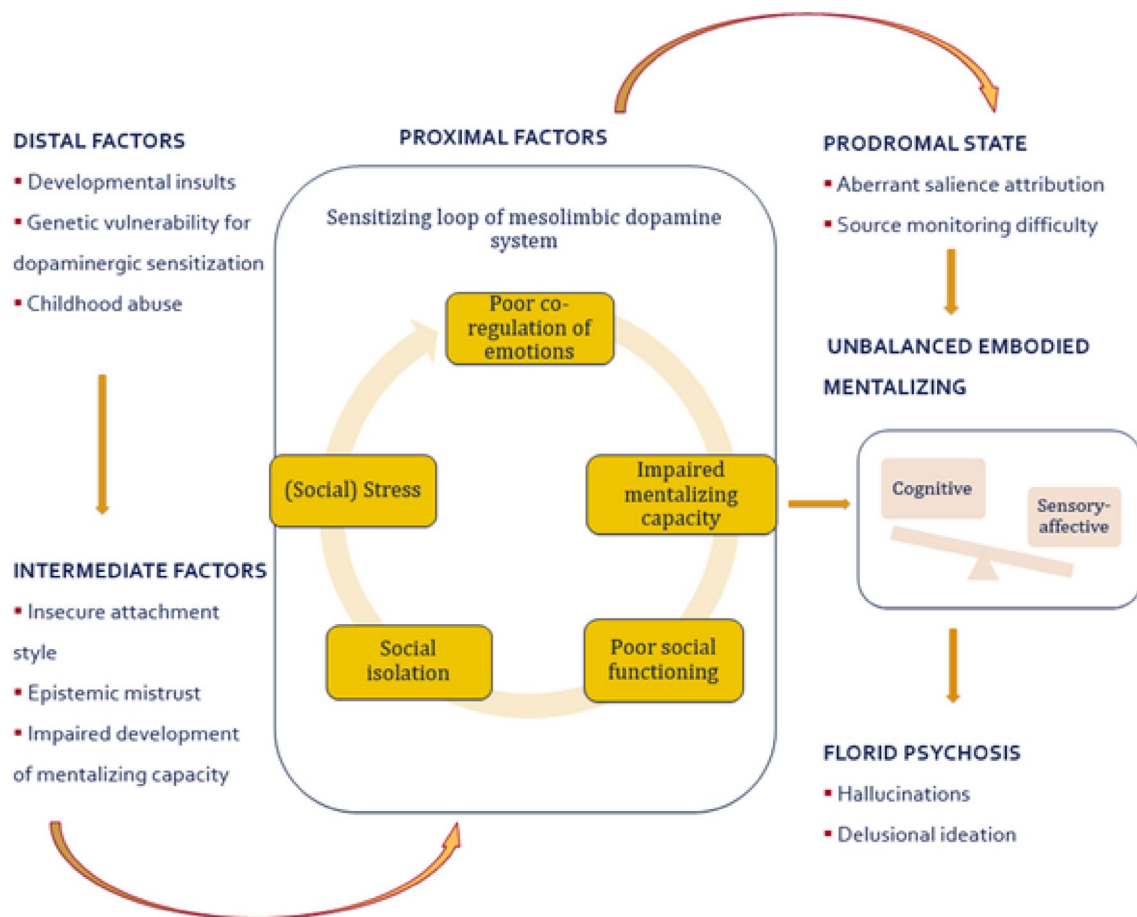
Fig. 1 A heuristic, mentalization based model of psychosis pathogenesis. Together with a constitutional liability caused by variant genes, developmental factors such as childhood abuse and insults, and intermediate factors such as insecure attachment and impaired mentalizing comprise a fertile ground for social difficulties to sensitize the mesolimbic dopamine system. Dismissing or disorganized attachment relationships to caregivers and peers may result in distress and poor co-regulation of emotions when in the proximity of others. This complicates social interaction, and may lead to isolation or withdrawal, and thus more stress and impaired mentalizing. The sensitization of the dopamine system in turn results in aberrant experiences. Due to epistemic mistrust, interpretation of these experiences may become progressively idiosyncratic and tenacious. If individuals attribute extreme certainty to sensory-affective experiences, this results in proper hallucinations. On the other hand, sensory-affective information or the viewpoint of others may not be used to update cognitive beliefs, resulting in delusional ideation

as resistance to other views when they conflict with pre-existing ones (Bentall and Swarbrick 2003) and a tendency for dogmatic beliefs (e.g., Mikulincer 1997). Please refer to Fig. 1 for a schematic overview of the model described in this section.

Part II: Implementation of Mentalization Based Treatment for NAPD

Epistemic hypervigilance in NAPD tends to complicate psychotherapeutic interventions because it entails that new perspectives that a therapist offers are seen as irrelevant, or attempts at coercion (Fonagy and Allison 2014). Secondly, many patients with NAPD experience difficulty mentalizing

under stress, meaning that psychotherapeutic interventions may be at risk of being too complex.

Principles of Mentalization Based Treatment

Mentalization based treatment (MBT) was specifically developed to address impaired mentalizing (Bateman and Fonagy 1999) and epistemic mistrust (Fonagy and Allison 2014), through four main principles: a 'not-knowing' therapeutic stance, unassuming interventions, a focus on currently felt affect, and careful adjustment to the patient's level of mentalization and arousal.

Firstly, MBT emphasizes that therapists adopt a 'not-knowing stance,' which entails that the therapist actively questions the patient and cultivates a genuine interest in

the patient's current experience without forcing any single interpretation upon the patient. When using this stance, the therapist communicates that she is trying to interpret the patient's actions in line with her own subjective experience and that the patient is being treated as an intentional agent (Debbané et al. 2016; Fonagy et al. 2017). Feeling "mentalized about" in therapy is thought to make the patient feel safe enough to think about himself in relation to his social world and how he operates in it (Fonagy and Allison 2014), which aids the restructuring of the "organization of thinking into less rigid, delusional and pervasive patterns of reality testing" (Pereira and Debbané 2018). Becoming more flexible in one's cognitive beliefs, opens up the opportunity to again learn from experience and other perspectives, which is thought to lead to an improvement of understanding one's self and the social world.

Secondly, interventions are kept unassuming and are aimed at getting the patient to identify and verbalize sensory-affective experiences. As Fonagy and Bateman (2006) underscore, "psychotherapists of many orientations often attempt to provide mentalistic understandings for issues that trigger intense emotional reactions at a time when the capacity for effective explicit mentalization is practically inaccessible." Therefore, MBT de-emphasizes the exploration of 'deep' unconscious motives, often present in psychodynamic treatment, instead focusing on more readily available subjective experience in the here and now.

Thirdly, the focus of both patient and therapist should be on the patient's current affective experience. By helping the patient actively reflect on their currently felt experience, MBT attempts to loosen the dominance of affect-driven modes of information processing on thought and behavior (Pereira and Debbané 2018) as verbalizing sensory-affective experience has been shown to downregulate the intensity of such affect at neural, physiological, and subjective levels (e.g., Torre and Lieberman 2018).

Fourthly, the intensity of interventions, that is, the complexity of mentalization required of the patient, is adjusted to the level of emotional arousal the patient is experiencing. In general, therapists attempt to keep questions and reflections at the maximum level of mentalizing that the patient is able to. Too little arousal provides too little material to actively reflect upon, while too much arousal will reduce a patient's ability to mentalize. At this level only supportive interventions should be uses such as affect validation.

Mentalization Based Treatment for Psychotic Disorder

Although MBT was originally designed for borderline personality disorder (BPD; Bateman and Fonagy 1999), given its focus on impaired mentalizing ability, it may be suitable for patients with NAPD as well. Recent articles have explored the applicability of MBT as a treatment for psychotic disorders or its prodromal states (Brent and Fonagy 2014; Weijers et al. 2016; Debbané et al. 2016) and the effectiveness of MBT for psychotic disorders is currently being examined in a randomized controlled trial (Weijers et al. 2016). However, disorder-specific characteristics do warrant some different approaches in MBT when dealing with patients with NAPD.

First, many NAPD patients tend to deactivate attachment-related affect to manage over-arousal triggered by social interaction. This may give a false impression of a lack of arousal. Moreover, negative symptoms such as flattened affective expression, can further complicate gauging the level of arousal of the patient and asking the patient to do complex mental work in the context of high arousal is not deemed helpful. MBT therapists must therefore be acutely aware on slight signs of emotional disturbance that are particular to NAPD, such as slowing of speech or decreased coherence.

Second, given their inclination for disorganized or dismissive attachment, NAPD patients do not tend to become very attached to their therapist or group members, at least in the initial phase of treatment. This means that attachment bonds tend to remain fragile for a longer time, and patients may find it relatively easy to drop out of treatment. Keeping patients 'in mind' through telephone calls or house visits following missed sessions, helps.

Third, we concur with Gumley and Liotti's (2008) observation that a more severe mentalization deficit distinguishes NAPD from BPD. Mentalizing in BPD seems to be characterized more by an instability rather than a deficit, while patients with NAPD tend to have a more structural difficulty detecting and interpreting sensory-affective experiences. Therapy sessions in the initial phase should therefore focus on elementary levels of mentalizing such as detecting affective experiences and verbalizing them.

Fourth, often antipsychotic medication has a substantial dampening effect on emotions. This may improve mentalizing ability in certain respects, but it can also make it difficult to talk about the emotional significance of events if such emotions are hardly experienced. It is therefore important that antipsychotic medication is attenuated with a specific focus on the ability to still feel affect.

MBT in Contrast to Other Psychotherapies for Psychosis

MBT differs from and shares similarities with other psychotherapies for psychotic disorders. Of them, cognitive behavioral therapy (CBT) for psychosis, has been the most studied and arguably most prominent intervention in the recent decades. Generally, the main goal of CBT is to correct negative beliefs and thoughts in order to treat symptoms. Both

CBT and MBT (e.g. Björgvinsson and Hart 2006) aim to increase the understanding of how cognitions and emotions affect behavior. However, CBT requires that some level of mentalization capacity is already present, asking patients to critically appraise their thoughts, whereas MBT does not. Furthermore, as a psychodynamic form of therapy, MBT is less directive, mainly aiming to help patients to better represent sensory-affective experience. Additionally, being rooted in attachment theory, MBT tends to be more directly concerned with the therapeutic relationship than CBT. Lastly, CBT takes a less 'holistic' approach to treatment, focusing on specific symptoms instead.

More akin to MBT, are the so called 'third wave' cognitive behavioral therapies. Whereas CBT was developed to treat specific symptoms, the third wave cognitive therapies, like MBT, have shifted their focus on the thinking process underlying symptoms rather than on the content of that thinking process (Lana et al. 2017). In other words, they have shifted their focus from *what* people think to *how* people think (Björgvinsson and Hart 2006). Social cognitive therapies for example tend to focus on specific aspects that underlie other-oriented mentalizing, such as affect recognition, social attribution biases or theory of mind (see Lana et al. 2017 for an overview of different approaches). Such approaches tend to differ from MBT, however, because they are more directive in nature and also tend to focus on just one distinct element of mentalizing. Additionally, such treatments do not target self-oriented aspects of mentalizing.

Perhaps most akin to MBT is Metacognitive Insight and Reflection Therapy (MERIT; de Jong et al. 2019). Ridenour et al. (2019) have previously noted that descriptions and case reports suggest that MBT and MERIT are largely compatible with one another. MERIT, like MBT, focuses on the identification of mental states, the differentiation of subjective experiences from objective reality, relating mental states to behavior, and integrating such knowledge into narratives. Similar to MBT, MERIT focuses on higher order cognitive processing by engaging the patient's ability to verbalize their subjective experience. The therapist takes an inquisitive stance towards the patient, searching for the patient's narrative rather than 'the truth'. Lastly, both emphasize that therapeutic questions should be adjusted towards the patient's level of metacognition.

There are however subtle distinctions between MERIT and MBT, which have been extensively detailed by Ridenour et al. (2019). Chief of these are that MBT specifically offers psycho-education prior to treatment to elaborate on mentalizing and its relation to attachment-processes and emotions. MBT has also been argued to pay more attention to the bodily aspect of emotional experience (Debbané et al. 2016). From the therapist perspective, MBT-training provides specific clinical guides to assess emotional arousal in sessions, as well as how to work safely within the client-therapist relationship (Brent 2009, Debbané et al. 2016). On the other hand, MERIT arguably focuses more explicitly on the patient's agency, by explicitly making the patient set the agenda for each treatment session. Lastly, group-therapy is an important element of MBT, but this is not the case for MERIT.

Clinical Vignette

To conclude, we will shortly illustrate the principles of MBT for psychotic disorder with a patient who suffers from paranoid delusions and finds it difficult to relate to others who do not share his beliefs.

Therapist: *Hi, how are you feeling?* (The therapist immediately focuses on current affect.)

Patient: *Yeah, I'm good; I've been on holiday with my girlfriend, which was mostly good*

T: *Mostly good?* (The therapist adopts a not-knowing stance)

P: *Yeah, we had fun, but I also had a falling-out with her, which is still bothering me. I tried to talk to her about my views on the world. I really wanted her to see my point of view*

T: *In what way is this still bothering you?* (Again, the therapist focuses on current affect.)

P: *I have experiences like this all the time, with friends, for example. I sometimes just feel that if they saw it my way, they would agree with me.* (Here, the patient is focused on the 'then and there,' but is not consciously attending his own current affective experience.)

T: *How does that make you feel now?* (Again focus is shifted to current affective experience.)

P: (The patient starts avoiding eye contact, and begins to talk in an agitated manner.) *I have put a lot of research into how the world works. We are being deceived; the evidence is there.* (The patient now shows signs of agitation; his answers do not clearly address the questions asked, and he withdraws from contact. In this state it is unlikely that his therapist will reach him. Being asked to think about his current feelings may have been too complex at this time and complexity needs to be scaled back at this point.)

T: *Hey, I see you are getting a bit agitated. This must be difficult for you!* (Here, the therapist tries to reduce the level of stress by validating the patient's feelings.)

P: (Shrugs) *It's no use talking about this stuff, because you people simply adhere to what we have been fed by the media, like sheep. I get*

the feeling that you are all against me, but I am used to it. (Here, the patient shows clear signs of impaired mentalizing. He has lost the ability to distinguish between individuals, lumping everyone in one category, he is now likely unable to consider therapist's perspective as meaningful.)

T: *I'm very sorry to hear that you perceive everyone is against you. That must be so difficult to bear, that perception that everyone is against you.* (Since the level of stress for the patient is too high, the therapist resorts to supportive interventions, validating his current affective state. By using terms like 'perceive' and 'perception' the therapist also addresses the fact that the patient's view is impressionistic, implying it does not necessarily reflect reality.)

P: (Relaxes visibly) *I don't necessarily think you are against me, but many people are.* (The ability to mentalize seems to return, as he regains the ability to distinguish between people.)

T: *Hey, it seems to me that you got a little worked up there, but that now you have also calmed down a bit. Could you help me understand what happened there?* (Here, the therapist tries to get the patient to reflect on the affective response he just had.)

P: *I get worked up about this stuff, and then I get angry with people for no reason*

T: *Shall we look at what happened there then? It seems like you lumped me together with people you dislike* (Here, the clinician consider the interaction that just occurred just as a prototype for interactions with others.)

P: *I react to people too quickly because I don't like it when people do not agree with me. I think I get upset, when I have the feeling people do not take me seriously.* (Here, through becoming conscious of his affective reaction to the therapist, the patient has developed some insight into the way he relates to others in general.)

Concluding Remarks

With this article, we attempted to show how mentalization theory may inform understanding and treatment of NAPD, especially regarding difficulties in mentalizing and epistemic mistrust. On-hand experience with mentalization based treatment for psychotic disorder has given the impression that, MBT may help loosen rigid interpretations of aberrant sensory-affective experience and thus improves social interaction.

References

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. N. (1978). *Patterns of attachment: A psychological study of the strange situation*. London: Psychology Press.

Aleman, A., & Larøi, F. (2008). *Hallucinations: The science of idiosyncratic perception*. Washington, DC: American Psychological Association.

Arnsten, A. F. (2009). Stress signalling pathways that impair prefrontal cortex structure and function. *Nature Reviews Neuroscience, 10*(6), 410.

Aydin, O., Balikci, K., Tas, C., Aydin, P. U., Danaci, A. E., Brüne, M., et al. (2016). The developmental origins of metacognitive deficits in schizophrenia. *Psychiatry Research, 245,* 15–21.

Baron-Cohen, S., Leslie, A. M., & Frith, U. (1986). Mechanical, behavioural and intentional understanding of picture stories in autistic children. *British Journal of Developmental Psychology, 4,* 113–125.

Bateman, A., & Fonagy, P. (1999). Effectiveness of partial hospitalization in the treatment of borderline personality disorder: A randomized controlled trial. *American Journal of Psychiatry, 156*(10), 1563–1569.

Bateman, A. W., & Fonagy, P. (2004). Mentalization-based treatment of BPD. *Journal of personality disorders*, *18*(1), 36–51.

Bentall, R. P., & Swarbrick, R. (2003). The best laid schemas of paranoid patients: Autonomy, sociotropy and need for closure.

Psychology and Psychotherapy: Theory, Research and Practice, *76*(2), 163–171.

Berry, K., Barrowclough, C., & Wearden, A. (2008). Attachment theory: A framework for understanding symptoms and interpersonal relationships in psychosis. *Behaviour Research and Therapy,* *46*(12), 1275–1282.

Birchwood, M., Connor, C., Lester, H., Patterson, P., Freemantle, N., Marshall, M., et al. (2013). Reducing duration of untreated psychosis: Care pathways to early intervention in psychosis services. *The British Journal of Psychiatry, 203*(1), 58–64.

Björgvinsson, T., & Hart, J. (2006). Cognitive behavioral therapy promotes mentalizing. In J. G. Allen & P. Fonagy (Eds.), *Handbook of mentalization-based treatment* (pp. 157–170). Chichester: Wiley.

Bowlby, J. (1980). *Attachment and loss* (Vol. 1). New York: Random House.

Brent, B. (2009). Mentalization-based psychodynamic psychotherapy for psychosis. *Journal of Clinical Psychology, 65*(8), 803–814.

Brent, B. K., & Fonagy, P. (2014). A mentalization-based treatment approach to disturbances of social understanding in schizophrenia. In *Social cognition and metacognition in schizophrenia* (pp. 245–259). Cambridge, MA: Academic Press.

Brookwell, M. L., Bentall, R. P., & Varese, F. (2013). Externalizing biases and hallucinations in source-monitoring, self-monitoring and signal detection studies: A meta-analytic review. *Psychological Medicine, 43*(12), 2465–2475.

Brunelin, J., d'Amato, T., Brun, P., Bediou, B., Kallel, L., Senn, M., et al. (2007). Impaired verbal source monitoring in schizophrenia: An intermediate trait vulnerability marker? *Schizophrenia Research, 89*(1–3), 287–292.

Choudhury, S., Blakemore, S. J., & Charman, T. (2006). Social cognitive development during adolescence. *Social Cognitive and Affective Neuroscience, 1*(3), 165–174.

Debbané, M., Salaminios, G., Luyten, P., Badoud, D., Armando, M., Solida Tozzi, A., et al. (2016). Attachment, neurobiology, and mentalizing along the psychosis continuum. *Frontiers in Human Neuroscience, 10,* 406.

de Jong, S., van Donkersgoed, R. J. M., Timmerman, M. E., Aan Het Rot, M., Wunderink, L., Arends, J., et al. (2019). Metacognitive reflection and insight therapy (MERIT) for patients with schizophrenia. *Psychological Medicine 49*(2), 303–313.

Eisenstadt, P., Monteiro, V. B., Diniz, M. J., & Chaves, A. C. (2012). Experience of recovery from a first-episode psychosis. *Early Intervention in Psychiatry, 6*(4), 476–480.

Fett, A. K. J., Viechtbauer, W., Penn, D. L., van Os, J., & Krabbendam, L. (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: A meta-analysis. *Neuroscience and Biobehavioral Reviews, 35*(3), 573–588.

Fonagy, P. (1989). On tolerating mental states. Theory of mind in borderline patients. *Bulletin of Anna Freud Centre, 12,* 91–115.

Fonagy, P., & Allison, E. (2014). The role of mentalizing and epistemic trust in the therapeutic relationship. *Psychotherapy, 51*(3), 372.

Fonagy, P., & Bateman, A. (2006). Progress in the treatment of borderline personality disorder. *The British Journal of Psychiatry,* *188*(1), 1–3.

Fonagy, P., & Bateman, A. W. (2016). Adversity, attachment, and mentalizing. *Comprehensive Psychiatry, 64,* 59–66.

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. (2017). What we have changed our minds about. Borderline personality disorder as a limitation of resilience. *Borderline Personality Disorder and Emotion Dysregulation, 4*(1), 11.

Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia.* London: Psychology Press.

Gallese, V. (2005). Embodied simulation: From neurons to phenomenal experience. *Phenomenology and the Cognitive Sciences, 4*(1), 23–48.

Green, M. F., Horan, W. P., & Lee, J. (2019). Nonsocial and social cognition in schizophrenia: Current evidence and future directions. *World Psychiatry, 18*(2), 146–161.

Gumley, A., & Liotti, G. (2008). An Attachment Perspective on Schizophrenia: The Role of Disorganized Attachment, Dissociation, and Mentalization. In Moskowitz, A., Schäfer, I. Martin, J. & Dorahy, M. J., (Eds.), *Psychosis, trauma and dissociation. Evolving Perspectives on severe psychopathology*. Wiley Online Library, pp. 97–116.

Gumley, A. I., Taylor, H. E. F., Schwannauer, M., & MacBeth, A. (2014). A systematic review of attachment and psychosis: Measurement, construct validity and outcomes. *Acta Psychiatrica Scandinavica, 129*(4), 257–274.

Harrington, L., Siegert, R., & McClure, J. (2005). Theory of mind in schizophrenia: A critical review. *Cognitive Neuropsychiatry, 10*(4), 24.

Hartwell, C. E. (1996). The schizophrenogenic mother concept in American psychiatry. *Psychiatry, 59*(3), 274–297.

Hasson-Ohayon, I., Goldzweig, G., Lavi-Rotenberg, A., Luther, L., & Lysaker, P. H. (2018). The centrality of cognitive symptoms and metacognition within the interacting network of symptoms, neurocognition, social cognition and metacognition in schizophrenia. *Schizophrenia Research, 202,* 260–266.

Hendryx, M., Green, C. A., & Perrin, N. A. (2009). Social support, activities, and recovery from seriousmental illness: STARS study findings. *Journal Behavioral Health Service Research, 36*(3), 320–329.

Holthausen, E. A., Wiersma, D., Cahn, W., Kahn, R. S., Dingemans, P. M., Schene, A. H., et al. (2007). Predictive value of cognition for different domains of outcome in recent-onset schizophrenia. *Psychiatry Research, 149,* 71–80.

Howes, O. D., & Murray, R. M. (2014). Schizophrenia: an integrated socio-developmental-cognitive model. *The Lancet, 383*(9929), 1677–1687.

Kantrowitz, J. T., Hoptman, M. J., Leitman, D. I., Silipo, G., & Javitt, D. C. (2014). The 5% difference: early sensory processing predicts sarcasm perception in schizophrenia and schizoaffective disorder. *Psychological Medicine, 44*(1), 25–36.

Kelleher, I., Keeley, H., Corcoran, P., Ramsay, H., Wasserman, C., Carli, V., et al. (2013). Childhood trauma and psychosis in a prospective cohort study: Cause, effect, and directionality. *American Journal of Psychiatry, 170*(7), 734–741.

Keysers, C., & Gazzola, V. (2007). Integrating simulation and theory of mind: From self to social cognition. *Trends in Cognitive Sciences, 11*(5), 194–196.

Kim, S. (2015). The mind in the making: Developmental and neurobiological origins of mentalizing. *Personality Disorders: Theory, Research, and Treatment, 6*(4), 356.

Kraan, T., Velthorst, E., Smit, F., de Haan, L., & van der Gaag, M. (2015). Trauma and recent life events in individuals at ultra-high risk for psychosis: Review and meta-analysis. *Schizophrenia Research, 161*(2–3), 143–149.

Lana, F., Africa Cruz, M., Sola Victor, P., Marti-Bonany, J. (2017). Social cognition based therapies for people with schizophrenia: Focus on metacognitive and mentalization approaches. *Schizophrenia Treatment.* SM Groups Open access ebooks.

Lane, J. D., & Harris, P. L. (2015). The roles of intuition and informants' expertise in children's epistemic trust. *Child Development, 86*(3), 919–926.

Leonhardt, B. L., Huling, K., Hamm, J. A., Roe, D., Hasson-Ohayon, I., McLeod, H. J., et al. (2017). Recovery and serious mental

illness: A review of current clinical and research paradigms and future directions. *Expert Review of Neurotherapeutics, 17*(11), 1117–1130.

Lind, S. E., & Bowler, D. M. (2009). Recognition memory, self-other source memory, and theory-of-mind in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 39*(9), 1231.

Lysaker, P. H., Buck, K. D., Fogley, R. L., Ringer, J., Harder, S., Hasson-Ohayon, I., et al. (2005). The mutual development of intersubjectivity and metacognitive capacity in the psychotherapy for persons with schizophrenia. *Journal of Contemporary Psychotherapy, 43*(2), 63–72.

Lysaker, P. H., & Dimaggio, G. (2014). Metacognitive capacities for reflection in schizophrenia: Implications for developing treatments. *Schizophrenia Bulletin, 40*(3), 487–491.

Main, M., & Solomon, J. (1990). Procedures for identifying infants as disorganized/disoriented during the Ainsworth Strange Situation. *Attachment in the Preschool Years: Theory, Research, and Intervention, 1,* 121–160.

McLeod, H. J., Gumley, A. I., MacBeth, A., Schwannauer, M., & Lysaker, P. H. (2014). Metacognitive functioning predicts positive and negative symptoms over 12 months in first episode psychosis. *Psychiatry Research, 54*(1), 109–115.

Mickelson, K. D., Kessler, R. C., & Shaver, P. R. (1997). Adult attachment in a nationally representative sample. *Journal of Personality and Social Psychology, 73*(5), 1092.

Mikulincer, M. (1997). Adult attachment style and information processing: Individual differences in curiosity and cognitive closure. *Journal of Personality and Social Psychology, 72*(5), 1217.

Mikulincer, M., & Shaver, P. R. (2012). An attachment perspective on psychopathology. *World Psychiatry, 11*(1), 11–15.

Nolte, T., Bolling, D. Z., Hudac, C., Fonagy, P., Mayes, L. C., & Pelphrey, K. A. (2013). Brain mechanisms underlying the impact of attachment-related stress on social cognition. *Frontiers in Human Neuroscience, 7,* 816.

Norman, R. M., Windell, D., Lynch, J., & Manchanda, R. (2013). Correlates of subjective recovery in an early intervention program for psychoses. *Early Intervention in Psychiatry, 7*(3), 278–284.

O'Driscoll, C., Laing, J., & Mason, O. (2014). Cognitive emotion regulation strategies, alexithymia and dissociation in schizophrenia, a review and meta-analysis. *Clinical Psychology Review, 34*(6), 482–495.

Owens, K. A., Haddock, G., & Berry, K. (2013). The role of the therapeutic alliance in the regulation of emotion in psychosis: An attachment perspective. *Clinical Psychology & Psychotherapy, 20*(6), 523–530.

Pereira, J. G., & Debbané, M. (2018). An integrative-relational approach in schizophrenia: From philosophical principles to mentalization-based practice. In I. Hipolito, J. Goncalves, & J. Pereira (Eds.), *Schizophrenia and common sense* (pp. 193–207). Cham: Springer.

Pinkham, A. E., Penn, D. L., Perkins, D. O., & Lieberman, J. (2003). Implications for the neural basis of social cognition for the study of schizophrenia. *American Journal of Psychiatry, 160*(5), 815–824.

Read, J., Fosse, R., Moskowitz, A., & Perry, B. (2014). The traumagenic neurodevelopmental model of psychosis revisited. *Neuropsychiatry, 4*(1), 65.

Ridenour, J., Knauss, D., & Hamm, J. A. (2019). Comparing metacognition and mentalization and their implications for psychotherapy for individuals with psychosis. *Journal of Contemporary Psychotherapy, 49*(2), 79–85.

Selten, J. P., van der Ven, E., Rutten, B. P., & Cantor-Graae, E. (2013). The social defeat hypothesis of schizophrenia: An update. *Schizophrenia Bulletin, 39*(6), 1180–1186.

Semerari, A., Carcione, A., Dimaggio, G., Falcone, M., Nicolò, G., Procacci, M., et al. (2003). How to evaluate metacognitive functioning in psychotherapy? The metacognition assessment scale and its applications. *Clinical Psychological Psychotherapy, 10*(4), 238–261.

Sprong, M., Schothorst, P., Vos, E., Hox, J., & Van Engeland, H. (2007). Theory of mind in schizophrenia: Meta-analysis. *The British Journal of Psychiatry, 191*(1), 5–13.

Sroufe, L. A. (2005). Attachment and development: A prospective, longitudinal study from birth to adulthood. *Attachment & human development, 7*(4), 349–367.

Stumbo, S. P., Yarborough, B. J. H., Paulson, R. I., & Green, C. A. (2015). The impact of adverse child and adult experiences on recovery from serious mental illness. *Psychiatric Rehabilitation Journal, 38*(4), 320.

Thomas, E. C., Muralidharan, A., Medoff, D., & Drapalski, A. L. (2016). Self-efficacy as a mediator of the relationship between social support and recovery in serious mental illness. *Psychiatric Rehabilitation Journal, 39*(4), 352.

Torre, J. B., & Lieberman, M. D. (2018). Putting feelings into words: Affect labeling as implicit emotion regulation. *Emotion Review, 10*(2), 116–124.

Trémeau, F. (2006). A review of emotion deficits in schizophrenia. *Dialogues in Clinical Neuroscience, 8*(1), 59.

Varese, F., Smeets, F., Drukker, M., Lieverse, R., Lataster, T., Viechtbauer, W., et al. (2012). Childhood adversities increase the risk of psychosis: a meta-analysis of patient-control, prospective-and cross-sectional cohort studies. *Schizophrenia Bulletin, 38*(4), 661–671.

Velthorst, E., Nieman, D. H., Linszen, D., Becker, H., de Haan, L., Dingemans, P. M., et al. (2010). Disability in people clinically at high risk of psychosis. *The British Journal of Psychiatry, 197*(4), 278–284.

Weijers, J., Fonagy, P., Eurelings-Bontekoe, E., Termorshuizen, F., Viechtbauer, W., & Selten, J. P. (2018). Mentalizing impairment as a mediator between reported childhood abuse and outcome in nonaffective psychotic disorder. *Psychiatry Research, 259,* 463–469.

Weijers, J., ten Kate, C., Eurelings-Bontekoe, E., Viechtbauer, W., Rampaart, R., Bateman, A., et al. (2016). Mentalization-based treatment for psychotic disorder: Protocol of a randomized controlled trial. *BMC Psychiatry, 16*(1), 191.

White, C., Stirling, J., Morris, J., Montague, L., Tantam, D., & Lewis, S. (2009). Predictors of 10-year outcome of first-episode psychosis. *Psychological Medicine, 39,* 1447–1456.