MBT Foundations

THE TALKING THERAPIES UK PRACTITIONER SERIES



MBT Foundations

Mentalising, attachment and the MBT stance

Foundation level · 7 hours of continuing professional development

TALKING THERAPIES UK

First edition · 2026

Published by Talking Therapies UK, talkingtherapies.co.uk.

Copyright © 2026 Talking Therapies UK. All rights reserved.

The moral right of the author has been asserted. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means without the prior written permission of the publisher, except for brief quotations in a review, and except that a purchaser of the accompanying course may print a single copy for their own study.

First edition, 2026.

This manual accompanies the MBT Foundations course in the Talking Therapies UK Digital Classroom, from which its content is drawn and against which it is kept current.

Revision 2. This printing was set on 1 November 2026 from the course as it stood that day.

A note on the clinical material. Every case in this book is a fictional composite written for teaching. No real person appears in it, however disguised. Nothing here constitutes clinical supervision, and nothing in it replaces the judgement of a qualified practitioner working within their competence, their professional body's code and their local governance. Where guidance is cited it was current at the date of this edition; readers must check the standing of any guidance before relying on it.

Set in Poppins and Roboto, the Talking Therapies UK typefaces. Figures drawn by the Talking Therapies UK teaching-diagram engine.

Contents

How to use this manual 1

Part 1 What mentalising is 3

1. Mentalising and its dimensions 4
2. When mentalising fails 9

How to use this manual

This comprehensive foundation course provides clinicians and trainees with a rigorous introduction to Mentalisation-Based Treatment, a modality rooted in the understanding that mentalising is a fundamental human capacity to interpret the behaviour of oneself and others in terms of mental states. You will explore the multidimensional nature of mentalising, examining how this capacity operates under optimal conditions and how it frequently breaks down or becomes distorted when faced with attachment stress or overwhelming emotions. The curriculum delves into the developmental origins of mentalising, illustrating how early attachment experiences shape our ability to make sense of the world and others. Building on this theoretical framework, the course elucidates the pivotal role of attachment theory within the MBT model, demonstrating how insecure attachment patterns can impede mentalising and lead to interpersonal difficulties. You will examine the robust evidence base supporting the efficacy of MBT, particularly in the treatment of Borderline Personality Disorder and other complex personality difficulties, gaining insight into the therapeutic outcomes associated with this integrative approach. The training clarifies the structural components of MBT, helping you understand how the treatment is organised to foster psychological change. Central to the MBT philosophy is the distinctive not-knowing therapeutic stance, a posture of curiosity and humility that allows the clinician to stay alongside the patient without premature interpretation. You will learn how to cultivate this stance, creating a safe mentalising space where the patient can explore their experiences. This course establishes the essential clinician attitudes required for MBT, ensuring you are well-prepared to progress to the advanced practitioner course on MBT technique. No prior experience with MBT is required, making this an ideal starting point for those seeking to deepen their understanding of mentalising and attachment.

This manual is the written form of a course taught in the Talking Therapies UK Digital Classroom. It is set at foundation level and carries 7 hours of continuing professional development for practitioners who complete the accompanying course and its assessment.

It is written to be read in order. Each part opens with what it is for; each

chapter teaches one idea to a working depth and closes with the points worth keeping and a practice task that converts reading into something you do. The tasks are the part most often skipped and the part that most reliably changes practice, and several of them are designed to be taken to supervision rather than completed alone.

The figures are teaching diagrams rather than decoration. Each is the conventional form of a model you will meet in the literature and in supervision, drawn in one visual language across the whole series so that the shape of an idea is recognisable wherever it reappears.

Where a chapter draws on published work, the reference is named in the text and gathered in the further reading at the back. Terms of art are defined in the glossary, which is worth reading straight through once before you begin.

What this manual will leave you able to do

- Define mentalizing and its four dimensions, and describe effective mentalizing in ordinary life
- Explain how attachment activation switches mentalizing off, and why that matters in therapy
- Recognise the non-mentalizing modes: psychic equivalence, teleological mode and pretend mode
- Describe how mentalizing failure produces the phenomena of personality difficulty
- Adopt the mentalizing stance: curious, not-knowing and actively enquiring
- Monitor your own mentalizing lapses in session and recover them

PART 1

What mentalising is

Mentalising, its dimensions, and how it fails under stress.

CHAPTER 1

Mentalising and its dimensions

Mentalising is the capacity this whole pathway is built around, and it deserves a definition precise enough to work with: mentalising is perceiving and interpreting behaviour, one's own and other people's, as driven by intentional mental states, by feelings, thoughts, wishes, beliefs and needs. It is what you are doing when you wonder why a friend's text sounded cold, when you notice your own irritation and ask what it is about, when you hold in mind that your colleague's snappishness might be their sick child rather than your failing. It is so ordinary that its presence is invisible, and so foundational that its absence explains half of clinical practice: without mentalising, behaviour, one's own included, arrives as brute event, uninterpreted and uninterpretable, and other people become objects, threats or instruments rather than minds.

The concept's power for clinicians comes from its dimensional structure, because mentalising is not one ability but a balance across four polarities, and imbalance, not absence, is the usual clinical finding. Self and other: mentalising can be aimed inward, what is happening in me, or outward, what is happening in them, and health is movement between the poles; the person locked in other-focus reads everyone's states while remaining a stranger to their own, the self-focused mirror image is fluent in their own turmoil and blind to its impact. Cognitive and affective: mental states can be known about, thought through, reasoned over, or felt with, resonated to, and each pole alone fails, the purely cognitive mentaliser producing accurate, bloodless accounts that touch nothing, hyper-affective mentalising drowning in resonance without comprehension, the felt storm mistaken for understanding. Automatic and controlled: most social cognition runs fast and implicit, effortless reading beneath awareness, and health includes the ability to slow down, to shift into deliberate, reflective mentalising when the automatic read misleads, precisely the shift that vanishes under stress. Internal and external: states can be inferred from interior evidence, what I know of her history and hopes, or read from exterior signals, the face, the posture, the tone, and over-reliance on the external, the anxious scanning of faces for verdicts, is a signature imbalance in the populations this pathway

serves.

Balanced mentalising, across all four dimensions, has a describable character, and clinicians learn to recognise it as a state rather than a trait: curiosity about mental states, one's own and others', held with the awareness that minds are opaque, that one's reads are fallible, that the other's inner world can surprise. The opacity point is central and easily missed: good mentalising is not accurate mind-reading, it is the humble, interested, revisable attempt, the knowing that one does not finally know. Its impostor, hypermentalising, the confident over-attribution of elaborate mental states on thin evidence, she looked at her phone, she is bored of me, she has been talking to my sister, they all think I am too much, is not excess health but a characteristic failure, common in adolescence and in borderline presentations, and distinguishing rich, tentative mentalising from florid, certain pseudo-mentalising is an assessment skill this course will train.

Why the capacity matters clinically can be stated as a chain. Mentalising is the basis of affect regulation: a feeling that can be identified, named and thought about is a feeling that can be modulated, which is why the capacity and emotional stability rise and fall together. It is the basis of relationships: intimacy is sustained mutual mentalising, conflict survived is conflict mentalised, and the interpersonal chaos of the personality disorder field is, in this model's reading, mentalising failure made chronic. It is the basis of the self: a coherent identity is a self that has been mentalised, experienced as continuous, intelligible, one's own, and the emptiness and diffusion this pathway's populations describe is what selfhood feels like when the mentalising that should constitute it is unreliable. And it is the basis of therapy itself, every school's: whatever else a treatment does, it asks the client to attend to mental states, which is why this model claims, plausibly, to name a common factor the brands share.

A useful discipline for the new practitioner is to run each clinical observation through all four dimensions in turn before settling on a formulation. Consider a patient who arrives late, sits down heavily and announces that the receptionist clearly despises her. On the automatic-controlled dimension she has reached a conclusion instantly, without any of the slower reflective work that might generate alternatives. On the self-other dimension her attention is fixed entirely on the other's supposed state, with no parallel curiosity about what she herself brought into the building. On the internal-external

dimension she has read hostile intent from external cues alone, a tightening of the receptionist's mouth, a delay in looking up. On the cognitive-affective dimension the affect is doing all the work; the belief feels true because the feeling is strong. None of these observations condemns her, and that is precisely the point of the exercise: dimensional analysis converts a provocative behaviour into a describable profile of imbalances, each of which suggests a direction of therapeutic travel rather than a verdict. The practitioner who habitually maps in this way also protects their own mentalising, because the task of locating a patient on four continua is incompatible with the certainty that fuels retaliation or despair.

Hold the dimensions against a person to make them clinical. A woman describes her partner's late return: he walked in, I knew he had been with someone, I could see it in his face, I did not need to ask, I threw his bag out of the window. Run the profile: external pole dominant, the face read as verdict; automatic dominant, no slowing, no checking; other-focused in content but with no genuine other-mind present, his perspective never consulted; affective flooding without cognitive balance; and the certainty, the absence of opacity, the hallmark. Nothing here is stupidity or malice; it is a mentalising profile under attachment stress, and it predicts exactly the treatment this pathway teaches: slow the automatic, recruit the cognitive, restore the question where the certainty was. Close by writing one such profile from your own practice: a single charged incident, mapped across the four dimensions, ending with the imbalance named. The next lesson teaches what happens when the capacity does not merely imbalance but collapses, into the prementalising modes that organise this model's clinical world.

Going deeper: the four dimensions at the bedside.

The four dimensions of mentalising earn their place when they stop being a diagram and start being a bedside instrument. Automatic–controlled is the master dial: ordinary social life runs on fast, intuitive reading, and health is the capacity to SHIFT to slow, explicit reflection when the fast read misfires — the capacity that collapses under arousal, which is why no one mentalises well mid-argument and why MBT manages temperature before content. Self–other names the balance failures: some minds read others acutely while their own states arrive as weather (alexithymic presentations), others are exquisitely self-aware while other minds stay opaque; personality difficulty often shows contagion, where the other's state floods the self

boundary. Internal–external distinguishes evidence bases — inferring from inner states versus from visible behaviour — and its imbalance produces both the paranoid certainty built on a glance and the naïve trust that ignores every glance. Cognitive–affective separates knowing about feelings from feeling them: hypercognitive mentalising (pretend-mode's engine) and affect-flooded knowing are opposite failures of the same integration.

The clinical use is locational: when an interaction derails, ask WHICH dimension tipped, and intervene on that axis — slow the automatic, borrow the missing perspective, point to external evidence, or warm the cognitive with 'and what does that feel like right now?'

A foundations exercise with lasting value: profile yourself under pressure. Every practitioner has a signature imbalance — most commonly certainty about the client's mind arriving fast and feeling like perception. Knowing one's own tilt is the beginning of the stance.

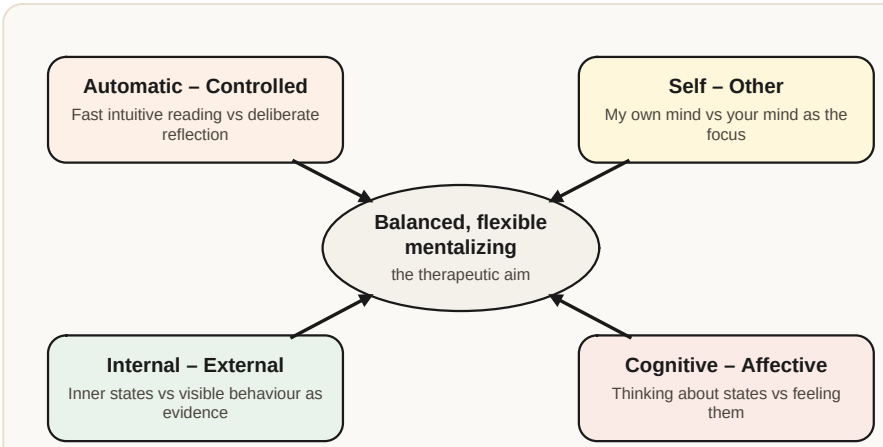


FIGURE 1.1 The four dimensions of mentalizing

Mentalizing runs on four balancing dimensions. Effective mentalizing moves flexibly along each; characteristic problems arise when someone is stuck at a pole, for example certainty about others' minds with no curiosity, or affect so hot that reflection collapses.

Points worth keeping

- Mentalising defined
- The four dimensions
- Balanced mentalising
- Why it matters clinically

PRACTICE TASK

Give an example of balanced and imbalanced mentalising.

CHAPTER 2

When mentalising fails

Mentalising is a regulatory capacity that enables individuals to interpret behaviour in terms of underlying mental states, both in themselves and in others, thereby allowing them to navigate the social world with flexibility and empathy. The foundations of this capacity were articulated by Peter Fonagy and Anthony Bateman, who described mentalising as a form of epistemic vigilance that balances the need for certainty with the openness to alternative perspectives. Central to their model is the notion that mentalising is not a static trait but a dynamic process that can be bolstered or undermined by contextual factors, most notably attachment stress. When an individual perceives a threat to the security of an attachment bond, the brain's survival circuitry is recruited, prioritising immediate emotional regulation over reflective functioning. This shift precipitates what Fonagy and Bateman term a "collapse" of mentalising, whereby the person regresses into one of three pre-mentalising modes: psychic equivalence, teleological mode, and pretend mode. Understanding the mechanisms of this collapse, and recognising the signatures of each mode, is essential for any therapist working within the mentalisation-based therapy (MBT) framework.

Attachment stress and collapse are best understood through the lens of Bowlian attachment theory, which posits that the internal working models formed in early relationships become the scaffolding for later affect regulation. When a caregiver is perceived as unavailable or rejecting, the infant's attachment system triggers a cascade of neurobiological responses, including activation of the amygdala and the hypothalamic-pituitary-adrenal (HPA) axis. In the adult, similar patterns emerge when relational threats are encountered – for example, a sudden silence from a partner, an unexpected criticism, or a perceived abandonment. The heightened arousal associated with such threats narrows attentional focus, reduces the capacity for perspective-taking, and drives the individual into a pre-mentalising state. In this state, the reflective function that normally mediates between affect and cognition is eclipsed, and the person becomes entrenched in a mode of thinking that is less nuanced and more concrete.

The first pre-mentalising mode, psychic equivalence, is characterised by a literal conflation of internal experience with external reality. In this mode, thoughts, feelings, and sensations are taken to be facts, and the distinction between “I think” and “I am” collapses. A client operating in psychic equivalence may assert, with absolute certainty, that “I am worthless” or “They hate me”, not as a hypothesis but as an incontrovertible truth. This mode is underpinned by a failure of mentalising to maintain the epistemic stance that mental states are representations that can be revised in light of new information. The rigidity of psychic equivalence often manifests as a lack of doubt, an inability to entertain alternative explanations, and a heightened sense of personal responsibility for events that are, in fact, beyond the individual’s control. Neurobiologically, the dominance of the limbic system over prefrontal regulatory regions during attachment stress reinforces this concrete, fact-like processing, limiting the capacity for abstract reasoning.

The second pre-mentalising mode, teleological mode, shifts the focus from mental states to observable outcomes, or goals. In teleological mode, the individual assumes that the world is best understood in terms of what is done rather than why it is done. The client may reason that “If I get the promotion, I will be happy” or “I must lie to protect my job, because the result matters more than the truth.” This mode reduces the complexity of interpersonal interaction to a simple cause-and-effect calculus, ignoring the internal experiences of self and other. It is often a defensive strategy employed when the affective load of considering another’s perspective feels overwhelming. By concentrating on tangible outcomes, the client sidesteps the anxiety that arises from uncertainty about others’ intentions. Teleological mode is closely linked to the concept of “goal-directed behaviour” as described by Fonagy and colleagues, wherein the mentalising system is bypassed in favour of a more primitive, action-oriented system that seeks to secure safety through concrete achievement.

The third pre-mentalising mode, pretend mode, involves a dissociative retreat into a fantasy or a role that protects the individual from the intensity of affect that would otherwise be felt. In pretend mode, the client may adopt a persona, a narrative, or a humorous stance that masks the underlying emotional pain. This is not merely “lying” in the colloquial sense, but rather a strategic disengagement from the reality of the present moment, allowing the individual to maintain a veneer of control while the affective storm is held

at bay. Pretend mode can be observed when a client jokes about a traumatic event, adopts an exaggeratedly confident demeanour, or describes themselves as “just a character in a play”. The protective function of this mode is reminiscent of the “false self” described by Winnicott, yet within the MBT framework it is conceptualised as a temporary suspension of mentalising that serves to preserve the self from overwhelming shame or guilt.

To illustrate how these modes operate in practice, consider the case of Maya, a 28-year-old woman referred for treatment following a recent breakup and a subsequent episode of self-harm. Maya’s attachment history includes intermittent caregiving from her mother, who was often emotionally unavailable due to work commitments. In the early sessions, Maya presents with a pervasive sense of worthlessness, repeatedly stating, “I am a failure” and “Nobody could ever love me”. The therapist notes that Maya’s statements are not presented as beliefs but as facts, indicating a collapse into psychic equivalence. When the therapist gently probes, asking “What evidence do you have for that belief?”, Maya responds with heightened anxiety and a refusal to consider alternative explanations, reinforcing the rigidity of the psychic equivalence mode.

As therapy progresses, Maya becomes increasingly preoccupied with the outcome of her recent breakup. She tells the therapist, “If I can just get my ex back, everything will be fine”. Here, Maya’s focus is on the tangible goal of reconciliation, rather than on the mental states that underlie the relational breakdown. She dismisses the ex-partner’s perspective, stating, “He must have left because he didn’t love me, so I need to prove he does”. This shift exemplifies teleological mode, where the pursuit of a concrete outcome (reunion) eclipses the exploration of underlying emotions such as abandonment, fear, or disappointment. The therapist, recognising this mode, reframes the conversation to invite Maya to consider the feelings that accompany the desire for reunion, thereby gently nudging her back towards reflective functioning.

In a later session, after a particularly distressing episode of self-harm, Maya arrives wearing a bright, flamboyant dress and adopts a joking tone, saying, “Well, I guess I’m just a drama queen, aren’t I?” She proceeds to recount the incident as if it were a theatrical performance, describing the act of cutting as “the climax of my latest show”. This narrative serves to distance her from

the raw shame and guilt associated with self-harm, illustrating pretend mode. The therapist acknowledges the humour while also inviting Maya to explore the underlying affect, saying, “It sounds like the performance is a way to keep the painful feelings at a distance – can we look together at what you might be feeling underneath the jokes?” By gently challenging the pretend mode, the therapist supports Maya in re-engaging with her internal emotional landscape, fostering the re-emergence of mentalising.

The therapeutic work of restoring mentalising after a collapse involves several interrelated strategies. First, the therapist must maintain a stance of curiosity and epistemic humility, signalling to the client that the therapeutic relationship is a safe space for exploring mental states without fear of judgment. This stance counters the defensive rigidity of psychic equivalence, where the client feels compelled to defend an immutable truth. Second, the therapist employs what Fonagy and Bateman term “mentalising-focused interventions”, such as the use of “mentalising-focused questions” that invite the client to consider alternative perspectives, for example, “What might your friend have been thinking when they didn’t reply?” Such interventions gently expand the client’s epistemic stance, encouraging the shift from certainty to possibility.

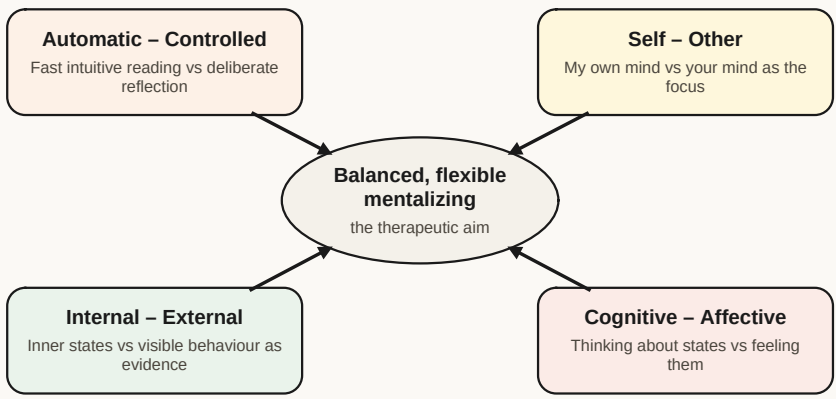
Third, the therapist attends to the affective regulation that underlies the collapse. When attachment stress is high, the client’s affect regulation system is overwhelmed, and the pre-mentalising modes serve as protective mechanisms. By providing a calm, attuned presence, the therapist helps to down-regulate the client’s arousal, thereby re-opening the prefrontal pathways necessary for reflective thinking. Techniques such as paced breathing, grounding exercises, and the use of “here-and-now” reflections are often employed to stabilise the client’s emotional state before attempting to explore mental states.

Fourth, the therapist must be vigilant for signs of the client slipping back into pre-mentalising modes during the session. In psychic equivalence, the client may present statements with absolute certainty; in teleological mode, they may reduce discussions to goal-oriented narratives; in pretend mode, they may adopt a detached or theatrical stance. Recognising these signatures allows the therapist to intervene promptly, either by gently challenging the rigidity, redirecting attention to mental states, or providing a containment for overwhelming affect.

Finally, the therapist supports the client in developing a more integrated sense of self that can tolerate ambivalence and uncertainty. This involves fostering a “mentalising stance” that acknowledges the provisional nature of mental representations, encouraging the client to adopt a stance of “I think that…” rather than “I know that…”. Over time, the client learns to hold multiple perspectives simultaneously, reducing the need for the protective pre-mentalising modes.

The evidence base for these interventions is robust. Randomised controlled trials of MBT for borderline personality disorder have demonstrated significant reductions in self-harm, improvements in interpersonal functioning, and increased capacity for mentalising as measured by the Reflective Functioning Scale. Moreover, neuroimaging studies have shown that successful MBT is associated with increased activation in the medial prefrontal cortex and decreased amygdala reactivity, suggesting a neurobiological shift toward greater regulatory control. The work of Allen, Fonagy, and Bateman (2008) further underscores the importance of attachment security in facilitating mentalising, highlighting that therapeutic relationships that provide consistent, responsive attunement can repair the attachment disruptions that precipitate mentalising collapse.

In summary, the collapse of mentalising under attachment stress is a central concept in MBT, manifesting through the three pre-mentalising modes of psychic equivalence, teleological mode, and pretend mode. Each mode reflects a distinct way in which the reflective function is overridden by more primitive, survival-focused processes. By cultivating a therapeutic environment that is both emotionally containing and intellectually curious, the therapist can help the client recognise when they have slipped into these modes, gently challenge the rigidity of their thinking, and support the re-emergence of a flexible, reflective stance. The case of Maya illustrates how these modes can appear in practice, how they can be identified, and how targeted interventions can facilitate the restoration of mentalising, ultimately leading to more adaptive interpersonal functioning and a reduction in self-destructive behaviours.



Balance on all four dimensions at once; being stuck at a pole is the difficulty. After Bateman and Fonagy.

FIGURE 2.1 The four dimensions of mentalizing

Mentalizing runs on four balancing dimensions. Effective mentalizing moves flexibly along each; characteristic problems arise when someone is stuck at a pole, for example certainty about others' minds with no curiosity, or affect so hot that reflection collapses.

Points worth keeping

- Attachment stress and collapse
- Psychic equivalence
- Teleological mode
- Pretend mode

PRACTICE TASK

Identify one prementalising mode in a recent interaction.