CHAPTER 3

How Does Mentalizing Develop?

Fonagy and colleagues' mentalization-based approach, like most psychodynamic approaches to psychotherapy, is rooted in early attachment relationships with caregivers, and is therefore fundamentally **developmental** in nature. This means we view every individual at any given moment as a complex product of her genetic endowment in interaction with early environmental (attachment) influences, and how those early forces interact, in turn, with more proximal (current) circumstances or stressors. The mentalization-based developmental model is rigorously grounded in empirical research (see Fonagy & Luyten, 2009, 2016; Fonagy & Sharp. 2008; Kim, 2015; Sharp & Kalpakci, 2015, for recent reviews). What we present below is a distilled version of this empirical research in order to provide an accessible understanding of the model for the development of typical mentalizing capacity.

Mentalization-based treatment is fundamentally developmental.

NORMATIVE TIMELINE FOR THE DEVELOPMENT OF MENTALIZING: THE PREMENTALIZING MODES

Table 3.1 maps out the normative timeline for the development of mentalizing as described in the developmental literature.

TABLE 3.1. Normative Development of Mentalizing

Developmental Period	Mentalizing Capacity
Infancy (0–6 months)	• Body awareness. • Some awareness of caregiver's attention. • Detects, responds to, and directs other's attention to her face and body. • Some awareness that she is separate from the world. Affectively rich communicative exchanges with caregiver.
Infancy (6–12 months)	• Physical self-recognition. • Begin to move from self-orientation to social orientation. • Begins to understand that actions have goals. • Development of **teleological thinking:** the use of observable physical reality to make inferences about goal-directed nature of actions (e.g., sees bottle of milk → drink bottle of milk). • Beginning to understand cause and effect. • Social referencing: seeks out the caregiver's emotional reactions to gauge her own affective reactions. For instance, a baby becomes fascinated with the way a branch is moving in the sunlight. The baby points at the branch and looks at Mother, waiting to see that her eyes focus on the branch. When they do, that is satisfaction enough that "she gets it!"
Toddlerhood (2–3 years)	• Understands that actions emanate from unobservable mental states but cannot actively reflect on mental states of others. • Implicit-automatic (System 1) mentalizing capacity (e.g., sister cries → comforts sister). • Can recognize basic emotion/facial expressions. • The idea of "me" emerges through pronoun use. • Autobiographical memory emerges. • Pretend play develops, which is the first practice in decoupling the mind from reality, but pretend games feel very real and child can get lost in magical thinking. • What is in my mind is real: **psychic equivalence**. • Begins to experience social emotions (shame, pride, embarrassment).
Early childhood (4–5 years)	• Monumental achievement in passing the false-belief task. • Can now decouple the mind from reality. • Achieve rudimentary explicit-controlled (System 2) mentalizing capacity: can verbally reason and interpret behavior and intentions behind behavior. • The capacity to deceive develops. • Begins making causal references to mental states. • Significant increase in mental state language. • Elaborated use of the pronouns *I*, *me*, *myself*.

(*continued*)

TABLE 3.1. (*continued*)

Developmental Period	Mentalizing Capacity
Middle childhood (6–11 years)	• Higher-order mentalizing develops ("he thinks she thinks that . . . "). • Begins to understand and differentiate between lies, jokes, irony, sarcasm. • Can carefully construct lies without leaking information that may lead to suspicion. • Significant increases in perspective-taking skills. • Understands social emotions (shame, pride, embarrassment). • Has the capacity for social exclusion, so creates "in-groups" and "out-groups" and may exclude or include others in play. • Understands personality traits and uses them to explain behavior. • Capacity for self-narrative begins to develop but is still rather concrete and unintegrated. • While mentalizing is in place, it is influenced by moral development and internalization of societal values—how one should think and feel. Mentalizing therefore lacks some authenticity and **pretend mode** is apparent in recounting of events and self narrative. • Capacity for social comparison emerges. • Self-evaluation largely positively skewed during early period of middle childhood and gradually becomes more realistic.
Adolescence (12–17 years)	• Social reorientation takes place: from parents to peers and romantic partners. • Identity consolidation begins. Identity defined as intrapersonal and interpersonal continuity; making sense/meaning from self-concept. • Cognitive development facilitates integration of self-representations and multiple self-hypotheses. • Autonomous self begins to emerge. • Cognitive development impedes control over emotions. Context sensitivity (shame, embarrassment, anxiety about social standing in relation to newly understood social rules) easily overwhelms cognitive mentalizing. • Shared reflection with peers. • Mature perspective-taking abilities, but self-consciousness and imaginary audience. • Stepping outside the social dyad to view self as a social object that is observed by others. • Abstract representations of self. • As control over emotions increases, greater sense of agency emerges.

Infancy (0–6 months)

Mentalizing capacity begins in the body, albeit in a reflexive, automatic way. A baby is endowed with the capacity to reflexively (automatically) know whether she is cold, hot, wet, dirty, or hurt. Because the baby cannot mentalize her bodily experiences with any explicit, cognitive reflective capacity, she is heavily reliant on her caregivers to help her mentalize her subjective experiences, which, as we said, mostly center on her bodily experiences. Her first job is to communicate this rudimentary subjective knowledge to her caregiver by crying or vocalizing. The following vignette illustrates how a mother might be responding to her crying baby:

> "Oh oh . . . hang on, little one . . . you need me to pick you up . . . let's see what's going on . . . let me peek into your diaper. . . . Oh, I see! You made a pee-pee. I think you're feeling all wet and uncomfortable. Let mommy help you . . . hang on . . . I'm putting you on your back and unzipping your onesie. Let mommy take off your diaper. There—there. Ah, now we get a wipe and clean you up nicely. Almost done. Just some paste and then I can wrap you back up. You'll feel nice and warm again."

We can imagine that during this interaction the mother is looking into her baby's eyes. She might be smiling gently, and her voice is soft and reassuring while she talks to her baby. The mother is busy *translating* or *mediating* the baby's subjective experience through the use of language, while at the same time making use of affect (emotions) to connect with her baby. In essence, the mother is mentalizing the baby's mind, even though the baby's mind is reflexively responding to her own bodily needs. The crucial point is that the mother is responding to the baby *as if* the baby has a mind (albeit a reflexive mind); and she uses affect and language to give meaning to the baby's subjective experiences—however rudimentary those experiences are from the adult's perspective (being hot, cold, hungry, hurt, wet or dirty).

The intensity with which the baby responds to her own discomfort or distress is determined by her **temperament**. Temperament is defined as the early-appearing variation in reaction and emotional reactivity and is first and foremost biologically (genetically) based.

Research has shown that babies with a difficult temperament will experience more distress in response to their bodily states—being wet, cold, warm, dirty, or hurt. They will express that intensity more loudly and for longer periods of time compared to their easy counterparts. They will be harder to soothe and will be more irregular and unpredictable in their bodily functions and patterns. Caregivers of difficult, temperament babies need *special skills* to help mediate their babies' subjective experience of their bodies.

Infancy (6–12 months)

During the second six months of life, the baby begins to move beyond her body and takes the first step to becoming social. This is a delightful time for parents, and the first social smile is an unforgettable moment of joy for most parents. The baby moves from self-orientation to social orientation and is able to direct her attention to her caregiver in more goal-directed ways. She is also able to direct her caregiver's attention to herself in a more goal-directed, less reflexive way, and she begins to make use of social referencing (that is, seeking out the caregiver's emotional reactions to gauge her own affective reactions). The caregiver continues to *connect the dots* for the baby by linking certain events or objects with one another. Below is another example of how an interaction may take place now that the baby is in the second 6 months of her life:

> "Good morning, little one! I heard you crying and came right over. How are you doing? Let me pick you up . . . oh, you feel nice and warm . . . are you hungry? Let's get your bottle ready . . . yes . . . it's breakfast time . . . see, it's light outside . . . come, let me hold you like this and we go get your bottle."

Here, we see how the baby is beginning to learn the connection between feeling hungry, getting her bottle, waking up, and breakfast time. She is beginning to understand cause and effect, albeit in a rudimentary way. This is the beginning of **teleological thinking** (a prementalizing mode)—a term we will return to many times throughout this book. Teleological thinking is defined as the use of observable physical reality to make inferences about the goal-directed nature of actions—we can think of teleological thinking as *concrete thinking*.

The use of observable physical reality to make inferences about the goal-directed nature of actions is based on a basic behavioral principle of stimulus and response. When I see my bottle, I'm going to get milk. When it is dark outside, it is sleep time. If I pee-pee in the potty, I'll get a star for my star chart. Even Donald Winnicott's *transitional objects*—the lovey blanket or soft toy that is used to repetitively tickle noses or to cuddle with—are teleological in that they are concrete, observable representations of attachment needs. Teleological thinking is regularly used throughout childhood to shape behavior based on stimulus–response pairings, and babies and small children (and our pets!) respond to these behavioral interventions early on.

Teleological thinking also shapes some of the earliest interactional patterns between infant and caregiver, shaping or forming the building blocks of mentalizing. The infant from early on can notice and respond to eye movements in other faces, and an early interactional pattern that many parents will recognize is when the baby learns to point at something that has caught her interest—the way the light is moving in some leaves overhead, for instance. The baby will point at the leaves that are so captivating, then pointedly look at Mom, then back at the leaves. Then checking back, the baby is satisfied to see Mom's eyes have moved to look for what is so special up there in the leaves. It's not sophisticated mentalizing, but at this tender age it will do: "I've seen this wondrous thing, but Mom hasn't got it yet. If Mom's eyes point at what I'm enjoying, then it means she is with me": the fact that Mom may still be wondering where she left the blanket, or when the car needs collecting from the repair shop, is neither here nor there—for the baby, if her eyes are up there with mine, we are connected, and all is good. It's the solution of a worry ("Perhaps I am all alone with this stimulating leafy sight!") by a simple physical outcome (Mom's eyes pointing at said foliage), but in it there are the earliest signs of interest in another mind.

Toddlerhood (2–3 years)

Whereas mental states are very much embedded in the body as well as physical objects during infancy, 2- to 3-year-old children begin to understand that actions may emanate from unobservable mental states. A 3-year-old may be saying something like "My sister is crying. She is sad." Thus, she is able to recognize basic emotions and

facial expressions. However, at this age, toddlers are not yet able to actively reflect on the mental states of others. The toddler is therefore still engaged mostly in automatic, implicit mentalizing without the option of explicit reflective mentalizing. But as the social circle of the toddler widens, the caregiver begins to play an increasingly important role in facilitating more explicit-controlled mentalizing. For instance, if a sibling is crying, a mother might say, "Oh no! Sissie is crying. What do you think might be going on? Should we ask her?" In other words, the caregiver is using language to expand mentalizing capacity not only in terms of self function (as in infancy), but also in terms of interpersonal function. The spurt in language development that occurs in toddlerhood is also associated with the emergence of *me* through pronoun use and the beginnings of an autobiographical memory.

An important milestone for mentalizing development during the toddler years is the development of pretend play, the first practice in decoupling the mind from reality, which, as discussed in Chapter 1, forms a key complement of the capacity to mentalize. However, pretend play still feels real for a child. Jean Piaget described this as "magical thinking"—the belief that one's thoughts, by themselves, can bring about effects in the world or that thinking something corresponds with doing it (e.g., "The soap slipped down the drain, so I can, too"). It is magical thinking that often complicates potty training—for instance, believing that it is possible to be flushed down the toilet with a poo-poo which came out of one's own body. It is magical thinking that causes fear of monsters in the closet, and it is magical thinking that also provides endless joy for toddlers dressing up as princesses or pirates. Once dressed, a 3-year-old becomes Anna or Elsa or Beauty or Peter Pan. In short, for the young child *what is in my mind is real*—and in mentalization-based terms, we call this **psychic equivalence** (another prementalizing mode we will return to again later).

Early Childhood (4–5 years)

Around age 4, we witness the monumental achievement of passing the false-belief task we referred to in Chapter 2. Children can now officially mind-read! That is, they can decouple the mind from reality and they can verbally reason and interpret behavior and intentions

behind behavior. In other words, we see for the first time the capacity for explicit-controlled mentalizing. With this, the capacity to deceive develops, although it is not very sophisticated, and most parents can see right through it.

During this period, there is also a marked increase in the use of mental state language. The parent's role continues to be that of mediator of subjective experience. To bring this role into sharper focus we can imagine the difference between a child playing with a stick in the water by herself versus playing with a stick in the water with an adult who engages the child in conversation about what she may find in the puddle, things that may be lurking under the rocks, and how the tide may be pushing the seawater in an out of the pool. We can equally imagine the difference between a small child crying alone after having his feelings hurt by a friend; vs. a small child crying in the arms of his dad feeling the warmth and strength of his father as he sobs. Put differently, caregivers continue to mediate the subjective experience of children to help create meaning for them. In this respect, Pnina Klein (1996), a developmental psychologist, wrote about the caregiver's role in creating "mediated learning experiences" for children. Elsewhere (Sharp et al., 2020; Sharp & Marais, 2022) we have described that a mediated learning experience begins with interactions on a preverbal level and is not related specifically to a modality, language, or content and is, therefore, a universal phenomenon. As a child matures, mediated learning experiences become verbal in addition to nonverbal and enable the child to benefit from experiences that he or she has not perceived directly, but can only perceive because an adult mediates them. The transmission of the past is made possible this way; and the awareness of the past and mediated anticipation of the future enable the child to expand his or her understanding of time and space. A child who receives mediated learning experiences develops a need for more mediation, that is, a need for events or objects to have meaning, a need to search for relations beyond the information provided by the senses at any given moment. In sum, mediated learning experiences enable further change of the individual through direct exposure to stimuli and allow a child to acquire basic structures that prepare her for future learning. Adult–child interactions can be considered as mediated learning experiences if they are intentional and reciprocal, if they transcend the satisfaction of an immediate need, and if they mediate meaning. We will again return to the idea of learning later

in this book, as learning forms an essential part of the mentalization-based approach to understanding psychological functioning. Learning from each other, passing on information, and creating the shared systems of meaning that make up culture is one of the key aspects that distinguishes humans from other mammals. Later, we will show how mentalizing acts as the key to an evolved *gating system* in the human mind, one that helps us in the crucial task of *deciding* whose minds we should open up our own to, and learn from, and to whom we keep that gate tight shut so that their influence on our mind is minimal.

Middle Childhood (6–11 years)

With basic ToM or mentalizing capacity in place, the middle childhood years see the development of higher-order mentalizing capacity. This includes what Happé (1994) referred to as second-order ToM ("He thinks she thinks that"), as well as lies, jokes, irony, and sarcasm. Parents endure hours of riddles as their 10-year-olds rejoice in the world of mind tricks and games. There is a sharp increase in perspective-taking skills such that the preadolescent is able to understand complex social interactions. The price these sophisticated mind-readers pay is, however, the sharp increase in social emotions (shame, pride, embarrassment, and so forth) and the risk of social exclusion, all of which develop with the increasing capacity for social comparison. We also begin to see a gender difference widen in social intelligence. Research shows that compared to boys, girls are already more advanced in mentalizing capacity at birth, and that this gap begins to widen in preadolescence and only normalizes again in adulthood when men catch up to women in social intelligence, albeit always lagging a little behind (Baron-Cohen, 2004).

An important development in middle childhood is the awareness that people have personalities. Children begin to use dispositional traits for the first time to explain others' behaviors and actions. They begin to develop a sense of their own personalities, albeit concrete. For instance, asked the question, "Can you describe yourself?" an 8-year-old may say, "I run really fast." Views on the self are also unintegrated and inconsistent. For instance, a 10-year-old may say she hates math on Monday, only to recant on Tuesday. Children at this stage show a positively skewed (unbalanced) self-concept and are

concerned with morality and how things *should be* and how things *should be done*. Pretend mode (the third prementalizing mode we will refer to many times throughout this book) characterizes this tendency to act in an *as-if* fashion based on expected rules or scripts rather than what naturally fits with the person or situation. Middle childhood is not the first time that pretend mode is detected in children. The toddler shows pretend mode, for instance, in her apparent competence when she says she is "a big girl" for using the potty. But it is during middle childhood when we see the lack of authenticity most clearly displayed given the strong push for what Erikson (1950) named "industry" (vs. inferiority). Erikson pointed out that the child at this stage of development feels the need to win approval by demonstrating specific competencies that are valued by society. Often, the enactment of a particular competence, even if it does not feel real, helps the preadolescent to avoid inferiority as she figures out her true talents and skills.

Adolescence (12–17 years)

Adolescence is another watershed period for the development of mentalizing capacity because it is during this stage that *adultlike mentalizing of the self* begins in earnest. Research has shown that during adolescence, two related and important changes occur, namely social reorientation and accompanying brain changes in the frontolimbic system (Guyer, Silk, & Nelson, 2016; Nelson, Leibenluft, McClure, & Pine, 2005; Sawyer, Azzopardi, Wickremarathne, & Patton, 2018). In tandem, these two processes facilitate increased perspective taking and mentalizing others, which in turn spawn tremendous growth in the capacity to mentalize the self. Let's unpack this complicated process a bit (see Sharp, Vanwoerden, & Wall, 2018, and Sharp & Wall, 2017, for a further discussion of this topic).

For some time we have known that adolescence is the developmental period during which children expand their attachment and social circles to go beyond parents to include peers and romantic partners. In this vein, the adolescent's primary influences shift from vertical (parents, teachers, and other key adults) to horizontal (their position in the social milieu). Middle and high school

In adolescence, adultlike mentalizing emerges.

generally bring with them a new social environment, which is for most adolescents a high-stakes environment on account of its potential for hosting episodes of social shame in ways that the primary (parental) environment simply did not. In that sense there is a strong sense of *life getting serious,* at least in terms of social standing. In addition, teachers and parents begin to hold more stringent expectations for adolescents. These environmental changes set the stage for new dimensions of social comparison (e.g., academics, extracurricular activities, appearance). Additionally, consequences of academic achievement gain greater weight with adolescents starting to consider possible future selves, such as college and potential occupations. At the same time, peer relations stabilize into intense and serious friendships while teens begin to navigate and balance these friendships with emerging romantic attachment relationships. This social reorientation and expansion coincide with expanded cognitive skills, specifically with regard to increased perspective-taking skills, as well as increases in self-consciousness and concern about the appraisal of others. However, there remains an underdeveloped prefrontal control over these new capacities. The end result in some cases is extreme or ineffectively applied use of these new cognitive capacities, such as overthinking or overreacting, which cannot be held in check by the as yet still weak inhibitory power of the prefrontal cortex.

The joint result of these two developments (social reorientation and uneven cognitive development) are, on the one hand, the capacity for shared reflection with peers such that one's personal goals become integrated with the goals of close others; and on the other hand, the development of an *imaginary audience,* referring to the perception that "others are as preoccupied with (and interested in) my behavior as I am," missing out on the fact that it is much more likely they are as preoccupied with *their* behavior, as I am with mine.

It has been suggested that the imaginary audience phenomenon is a function of the separation–individuation process of adolescence. Constructing an imaginary audience creates a sense of closeness and importance among peers as adolescents renegotiate relationships with parents, which reflects the expansion of intimacy and close relationships beyond the family system to peers and potential romantic partners. For the first time, adolescents are able to step outside the (parent–child, teacher–child) dyad and view themselves as social objects (from the outside in). This is a complex process because they have to

integrate multiple self-hypotheses with feedback from peers, parents, teachers, and the wider social environment. They have to reflect on themselves in relation to others to decide which perspectives to internalize as defining features of their identity. Moreover, whereas preadolescents' unreflective and positively biased self-acceptance buffered them from potential negative self-images, adolescents begin to form more realistic views based on multiple perspectives on themselves. Whereas cognitive constraints on self-reflection enabled preadolescent children to compartmentalize different aspects of the self, adolescents no longer have that luxury and have to begin the hard labor of forming a consolidated identity. This has the potential of leading to doubt and uncertainty as adolescents consider multiple perspectives and opinions. Self-representations may become removed from concrete, behavioral evidence and therefore may be inaccurate. This means that adolescence is for some individuals a time of increased intrapersonal conflict, confusion, distress, and potential instability in self-representation. Conversely, it is this potential for uncertainty (however vulnerable it may feel) that opens up the gateway to mature mentalizing capacity. Adolescence is therefore the quintessential double-edged sword: with maturation of social-cognitive capacity, vulnerability for derailment increases—therefore it requires quite a bit of scaffolding from the social environment to help adolescents get over this important developmental hump.

It is not until late adolescence that an adolescent is able to integrate various self-representations to resolve apparent contradictions. By then, the ability to develop narratives that explain how chronological events in her life are linked comes online (this is referred to as **causal coherence**). Additionally, by mid- to late adolescence, individuals are able to identify overarching themes, values, or principles that integrate different events in their life, called **thematic coherence**. Both causal and thematic coherence becomes possible thanks to the adolescent's newly acquired ability for higher-order abstraction, which is used to meaningfully integrate what previously seem contradictions in her self-representations and allows for the individual's identity to consolidate. Dan McAdams (2015) talks about the "binding of personality" to capture the coming together of disparate aspects of personality functioning into a coherent whole. We have referred to this as the developmental period where self and interpersonal functioning (recall LPF or Criterion A, which we discussed in Chapter 1) bind

into a unidimensional severity criterion (Sharp & Wall, 2021). This achievement is extraordinary. For instance, a young person may be able to recognize that she talks rebellion with their peers but acts as a dutiful loving grandchild with her elderly grandmother, and instead of being burdened by feeling like a hypocrite or fake, she can construct this as an example of her kindness and adaptation to the fact that Grandma is unlikely to change at her age. Indeed, by late adolescence, individuals start to normalize potential contradictions in self-representations, which serves to reduce internal conflict. As adolescents move into young adulthood, they gain a greater sense of **agency** as they take steps to become their future selves. Of course, for some adolescents this process is not as smooth as for others. We will return to how personality development (binding) may go awry in Chapter 4.

THE ATTACHMENT ROOTS OF THE CAPACITY TO MENTALIZE

Throughout the discussion on the normative developmental progression of mentalizing capacity, we have mentioned the role of the caregiver in this progression. In this section, we further expand on that role.

It is universally accepted that sensitive early caregiving leads to positive cognitive and socioemotional outcomes for children. John Bowlby's (1973, 1980) attachment theory suggests that the caregiver's capacity to sensitively respond to a child's physical and emotional needs is important for the development of secure internal working models of the self and others. Internal working models are *representations* of attachment relationships and become the blueprint (or cognitive schema) that individuals use to understand and manage their attachment relationships as they mature through childhood, adolescence, and adulthood. A secure attachment representation means that an individual trusts that an attachment figure (e.g., a parent, husband, wife, best friend) is available to them and will meet their emotional and physical needs. In contrast, an insecure attachment representation is characterized by mistrust and expresses itself by either dismissing the attachment figure (or the possibility of needing them), or being preoccupied with whether the attachment figure actually cares about them.

Research over the last 60 years has shown that secure attachment

is associated with resilience and positive outcomes in children and adults while insecure attachment is associated with negative outcomes. For instance, a recent meta-analytic study of nearly 6,000 children confirmed that children with a secure attachment in the early years are significantly less likely to develop behavior problems across childhood (Fearon, Bakermans-Kranenburg, van IJzendoorn, Lapsley, & Roisman, 2010). Early security has been associated with lower rates of delinquent behavior and more positive peer interactions in adolescence. In contrast, attachment insecurity has been shown to be associated with suicide-related behaviors, greater use of residential treatment and inpatient admissions, and a range of psychopathology including internalizing and externalizing problems. Research has also clearly documented long-lasting and severe psychological problems as a result of maltreatment across biological and psychological domains, which exemplifies arguably the most toxic disruption of the early caregiving environment (Cicchetti & Toth, 2005).

According to mentalization-based theory, secure attachment is enabled via the mechanism of caregiver mentalizing. In Chapter 2, we introduced the concept of mentalizing, so by now you should have a fairly good understanding of what is meant by it. Parental mentalizing refers to mentalizing specifically within the caregiver–child attachment context. It is defined as both a cognitive process, akin to psychological insight or perspective taking, and an emotional process, that is, the capacity to hold, regulate, and fully experience one's own and the child's emotions in a nondefensive way without becoming overwhelmed or shutting down (Slade, 2005). In the attachment literature, parental mentalization is also referred to as **parental reflective function** (Fonagy, Gergely, Jurist, & Target, 2002) to describe the parent's capacity to reflect upon both her own and/or the child's internal mental experiences within the parent–child relationship as they manifest in parental descriptions of the ongoing, current, and evolving relationship to the child. Parental reflective function can be classified and rated using a standardized coding scheme based on attachment interviews or interviews specifically designed for assessing parental reflective function. If these descriptions of the relationship with the child show (1) an awareness of the nature of mental states, (2) an explicit effort to tease out mental states underlying behavior, and (3) recognition of the developmental aspects of mental states (and mental states in relation to the interviewer), then

the caregiver is rated as high on reflective function/mentalizing. Put differently, if the caregiver's descriptions of the child contain reference to mental states as described above, then the caregiver is treating the child as a psychological agent, that is, someone with a mind. And in acknowledging that the child has a mind, the caregiver is also acknowledging the child's *autonomy* and agency—that she is a person in her own right.

Recall the nonmentalizing homework example discussed in Chapter 2 (see also Sharp et al., 2020). This example demonstrates a caregiver's challenge in mentalizing her child when her own internal resources are low. When internal resources and assets are limited, it is common for caregivers to take *shortcuts,* which potentially disregards the child's agentic self (her autonomy). In that example, the fastest way in which Mom imagines solving the homework problem is by *making* her daughter get up and sit down to do it. In the mentalizing framework, the mother would be described as nonmentalizing and functioning in the *teleological mode,* which we referenced as a prementalizing mode that is normative in small children, and which we will discuss in more detail in Chapter 4. The mother's desire to get the homework done comes from a noble source—she has good intentions! She wants her daughter to do well at school because she knows that a good education will facilitate a bright future. Perhaps Mom did not have similar opportunities growing up and it's hard for her to see her daughter potentially squandering the education that mom never had. The end result of taking the shortcut, however, is that goals are not accomplished (the homework is still not done), her daughter is upset and crying, and Mom feels even more emotionally depleted. Here we see how less sophisticated **prementalizing modes** of thinking are common in we adults too! Sometimes teleology directs us to just the right thing to do (you feed a baby when she is hungry, you change a diaper when it is full, you grab a child if she is falling), but sometimes we switch into it a little faster than is helpful, creating more conflict or misunderstanding than might have been the case if there had been room for some mentalizing beforehand.

As we describe elsewhere (Sharp et al., 2020), the Mentalizing Stance provides the parent with an alternative to the shortcut, the idea being that when a parent is in a Mentalizing Stance, she is able to **slow down** the interaction and move herself and the child to a more reflective, uncertain, curious, and fluid mode. In other words, by keeping

in mind where the child is mentally in the moment, the parent is able to slow down the interaction and treat the child as a psychological agent with thoughts, feelings, needs, and desires different from her own. It is through this kind of parental mentalization that mentalizing capacity, autonomy, and self-regulation are fostered in the child because the child's mind is *minded*. The child whose mind is minded in this way can use the mind of the carer as a kind of mirror—one that is far more informative than a glassy version, in that she receives dynamic information on how she is perceived by another (trusted) mind, not just how the light bounces between her mind and a piece of glass. If you now revisit the mentalizing homework example provided in Chapter 2, you can see the difference in how the mother is minding the child's mind, resulting in self-regulation and positive affect instead of dysregulation and negative affect exemplified in the nonmentalizing example. Importantly, when mentalizing takes place, interaction partners feel *connected* and feel a *sharing of joy*.

Empirical research supports the importance of parental mentalizing for socioemotional development. For instance, in a now-classic study by Fonagy and colleagues (Fonagy, Steele, & Steele, 1991), *prenatal* parental reflective function was shown to be predictive of *subsequent* attachment security of the infant and of the mentalizing capacity (ToM performance) of the same children during the preschool years, and considerably more so than the parent's prenatal attachment style assessed by the Adult Attachment Interview. In another study, we also demonstrated that low maternal accuracy in predicting children's mentalizing responses is associated with ineffective mentalizing in the child (Sharp, Fonagy, & Goodyer, 2006). In this study, we also showed that superior maternal mentalizing is not associated with better child outcomes compared to *good-enough maternal mentalizing*. In other words, parents do not have to be in a Mentalizing Stance all of the time to ensure positive outcomes in their children. A little above average would suffice! In fact, we would argue that breakdown in mentalizing provides the stimulus for the child to practice mentalizing of the other, as well as outstanding opportunities for the parent to practice and model mentalizing with their child. When a parent fails to make sense of her child in a way that is comfortingly contingent to

The parent can slow down the interaction and treat the child as a psychological agent.

the child's experience, it may be just that mild discomfort that stimulates the child into wondering (mentalizing the other): "Why is that? Oh, Mom looks a bit worried about the car/dog/neighbor, perhaps she just didn't notice this time . . . " If we now consider, for instance, the nonmentalizing homework example we provided in Chapter 2, we can see how a breakdown in mentalizing provides a great opportunity for Mom to make a **mentalizing comeback**.

Box 3.1. A Typical Late Afternoon: Mentalizing Comeback

A mother, arriving tired at home after a full day of work, finds her 8-year-old daughter has not completed her homework as previously agreed upon. The mother puts down her bag, sighs, and looks at her daughter, who is sitting in front of the television watching a favorite show. "What?" says her daughter. Mom responds by saying, "You know what." Her daughter appears baffled. Mom sighs again and reminds her daughter in a somewhat exasperated tone that they agreed at school drop-off that the 8-year-old would complete her homework at after-school care. Her daughter explains that she forgot, and Mom says, "Well, that's not good enough. Go sit down now and do your homework while I start dinner. No buts! Now! Go sit down." The daughter becomes distressed and says that she wants to finish her show. Mom becomes more exasperated and says, "I don't want to be saddled with your homework after dinner. Do it now! Or no more television for you for the rest of the week." By now, the daughter is crying and runs off to her room.

Ten minutes later, Mom goes up to her daughter's bedroom, sits next to her, and takes her daughter's hand. "Sweetheart, I want to apologize. When I came in from work I was tired and impatient. I did not notice that you were in the middle of your favorite TV show and I interrupted you. I wanted to get the dinner done and I panicked. I think I came over a bit harsh. Did it feel that way to you?" Her daughter replies, "Yes, Mommy. I got really upset." Mom says: "Yes, I could see that. And I'm sorry. Can we start over? What would have been a better way for me to bring up that the homework was not done?" Daughter replies, "Hmmm, I don't know. Maybe just pause my show first and ask me whether we can talk?" Mom says, "Yes! You're right . . . let's practice that and see where it goes."

The *maternal mind-mindedness* displayed in Box 3.1 was operationalized by Meins and colleagues who demonstrated in a series

of empirical studies (Meins, 1997; Meins, Fernyhough, Fradley, & Tuckey, 2001) an association between secure child attachment and referential tendencies in infant language acquisition, perspective taking during pretend play, and mentalizing capacity as evidenced by passing a ToM task at age 4. In addition, mothers of securely attached children presented their children with information and instructions that were comprehensible and pitched within the child's zone of proximal development. Such mothers also used speech that contained more mental state terms when describing their children. The capacity of the parent to keep the child's mind in mind during interaction is therefore a powerful predictor of attachment, quality of interactions, and the child's own developing mentalizing capacity.

The parent's capacity to keep the child's mind in mind is powerful.

FROM ATTACHMENT TO LEARNING

A central mechanism described in mentalizing theory and research by which the caregiver achieves the Mentalizing Stance with her baby, is referred to as **affect mirroring**. This refers to the attachment figure's ability to respond with *contingent,* **marked** and *ostensive* affective displays of her own experience in response to her infant's subjective experience, which in turn makes possible the child's development of coherent second-order representations of these subjective experience (Fonagy & Luyten, 2016; Kim, 2015). Efrain Bleiberg has often recounted a scenario in which a baby cries profusely, her little face red and all scrunched up. In response, the mother may look into the eyes of her baby and with a slightly cartoonish imitation of a cross face, but speaking in a reassuringly singsong tone of voice, say, "Oh Sara! What's up little one? You're so upset! Come here, let me hold you just a bit." In this example we may say the mother's communication is marked because she demonstrates that she understands the infant's internal state (her upset face and the sense of urgency this conveys), while concurrently signaling that her own upset facial expression concerns *what she thinks might be the case for* the infant, not her own state of mind. She achieves this by modifying (e.g., exaggerating, slowing down) her display of the infant's affect, rendering

it perceptually distinguishable from her expression of her own affect. She is in that sense showing to her baby her best estimation of what she *thinks* is the baby's state of mind, but is also clearly separating her own mental state from that of the baby. And of course, if all is basically going okay, she is doing all this in real time, and with hardly any conscious effort. Her communication can also be described as rich in *ostensive communicative cues* (Csibra & Gergely, 2009) because she makes direct eye contact with the infant, slightly tilting her head toward the baby, speaking with a singsong "motherese" intonation, and calling the infant by name. In fact, the word *ostensive* literally translates as "reaching out" or "pointing." When we, as humans, engage in ostensive cueing, we are signaling to the recipient at the other end of the interaction that we are about to share some important information—that we are about to "teach" the person something. In this example, through her marked, contingent affect mirroring, the mother is communicating to the baby that she recognizes the baby's distress, that it makes Mom feel sad for her, but that Mom is able to help regulate the distress without becoming distressed herself. Mom is showing Sara that she can shoulder Sara's distress and that Sara and Mom are two separate people—with their own respective minds that contain unique subjective perspectives on what is happening. George Gergely has described this process as the mother demonstrating **pedagogical intention** (intention to teach), signaling to the infant that her expression or utterances concerns the infant and what unfolds within the infant, separate from herself. Over time, through regular exposure to this kind of contingently matched affect mirroring that gradually internalizes, the child first develops an awareness of her subjective internal state, which sets the stage for increasing self-awareness, and increasing control of internal states—in other words, self-regulation.

That's what happens when all's going well. Contrast this process with a scenario where Mom is not able to respond in this marked, contingent fashion, but instead starts crying to the same level and intensity as her infant—perhaps out of exhaustion or exasperation. In this scenario, the communication is *not* marked, but instead offers a perfect mirroring of the baby's distress. The mother tragically fails to communicate that she can shoulder the baby's distress. Her mind and the mind of the baby remain undifferentiated, and the baby learns

neither self-regulation, nor individuation. Instead, she learns that her distress (her affect; her mind) has a powerful and negative effect on her mom and that her mind is potentially dangerous. She learns to either not share her subjective experience with her mom (associated with a dismissing attachment style) or to up the ante by increasing the intensity or duration of her protest (associated with a preoccupied attachment style). We will return to how mentalizing development goes awry in Chapter 4.

EPISTEMIC TRUST

In an extension of the mentalization-based theory, Fonagy and colleagues have introduced the construct of epistemic trust (Allison & Fonagy, 2016; Bo, Sharp, Fonagy, & Kongerslev, 2015; Fonagy & Allison, 2014; Fonagy & Luyten, 2016). Epistemic trust is defined as "an individual's willingness to consider communication conveying the knowledge from someone as trustworthy, generalizable and relevant to the self" (Fonagy & Luyten, 2016) or the ability to appraise incoming information from the social world as accurate, reliable, and personally relevant because it holds broad social value rather than simply personal value to its original bearer. Appraising it this way will allow for that information to be incorporated into the learner's existing knowledge domains (Fonagy, Luyten, & Allison, 2015; Sperber et al., 2010). As we have discussed, in secure attachment relationships, parents consistently adopt a Mentalizing Stance toward the child, seeing the child as an intentional psychological agent with a mind and attempting to make sense of the child's behavior as arising from underlying mental states. The caregiver conveys understanding of the child's subjective experience in a way that is accurate (i.e., personally relevant) and marked as the parent's *representation* of the child's mental state. Marked communication serves as an ostensive cue that signals to the child that socially generalizable and personally relevant information is being communicated, effectively inviting the child to suspend epistemic vigilance to make use of helpful social information (Fonagy & Allison, 2014; Fonagy et al., 2015). When the caregiver mentalizes the child, it opens the *epistemic (or learning) highway* for the child. It communicates to the child that the

information the parent is conveying is important, relevant to them, and helpful. In a state of epistemic trust, the child is therefore able to accept culturally transmitted knowledge from the caregiver. The child is learning!

As it turns out, *learning is rewarding*. Brain research has shown dopamine increases when a person perceives stimuli that predict rewards. And dopamine feels good. Dopamine is what is released when people inject drugs, when they gamble, when they make love, and when they win (Ross, Sharp, Vuchinich, & Spurrett, 2008). These dopamine spikes are a dominant mechanism of reward learning within the brain (hence their addictive power). Recall our example earlier in this chapter when we recounted how a mother talks to her baby after her baby just woke up. Mom says: "Let me pick you up . . . oh, you feel nice and warm . . . are you hungry? Let's get your bottle ready . . . yes . . . it's breakfast time . . . see, it's light outside . . . come, let me hold you like this and we go get your bottle." In this example, the baby is pairing morning time with her bottle—the latter being rewarding (not to mention Mom's warmth when being picked up and cuddled). Over time, her brain begins to *expect* her bottle when it is morning time—just like Pavlov's dogs salivate when they hear a bell ring. We call this reward expectancy. She has learned to expect her bottle, and neuroscience tells us that every time we expect a reward, dopamine fires. It therefore makes sense that some neuroscientists have called dopamine a "learning signal" (Glimcher, 2003).

This simple example illustrates how a baby learns that morning time has something to do with breakfast time. It is important to note, however, that the baby cannot learn this information alone. Her mother teaches her this information by *mediating* the baby's subjective experience (hunger). In essence, the mother is mentalizing the baby's internal states. If she gets it right, what the baby learns is important (I'm hungry!), relevant (it's me who is hungry, not my mom!), and helpful (if I go with my mom I will be fed). To repeat, the knowledge gained from the interaction with Mom is important, relevant, and helpful. In turn, because this knowledge, over time, turns out to be important, relevant and helpful to the child, and because dopamine fires while she learns, *the child learns that learning is rewarding,* and begins to seek out learning herself. Pnina Klein and colleagues suggest that the child's **needs system** is stimulated

through this kind of mediated learning (Klein, 1996; Sharp et al., 2020; Sharp & Marais, 2022). What that means is that the child develops a *need* to (1) seek clarity of perception, (2) search meaning and excitement, (3) have successful experiences and complete tasks, (4) seek information beyond sensory experiences, explore, and ask adults for help, and (5) think before doing. In sum, the child becomes an *agentic learner* who can make use of the environment outside the home (school, peers, extracurricular activities, and so on) to further learn how to live effectively and happily.

With the introduction of the epistemic trust concept, Fonagy and colleagues suggest that *mentalizing is closely connected to learning.* Without consistent and sensitive caregiving, individuals may remain insulated from important learning experiences, which contributes to the cognitive rigidity that is one of the hallmarks of several forms of psychopathology, in particular personality pathology. It is to this topic that we turn to next in Chapter 4: difficulties in mentalizing—and ultimately, difficulties in learning from experience, as well as from the social environment.

CHAPTER SUMMARY

In this chapter, we aimed to increase your understanding of the developmental origins of mentalizing capacity. Doing so is important because in order to understand how mentalizing development goes awry in psychopathology, we must first understand its typical (normative) development. As we have shown, the capacity to mentalize is the culmination of several developmental prementalizing steps toward mature mentalizing. The prementalizing modes that precede the onset of mature, adultlike mentalizing are teleological mode, psychic equivalence, and pretend mode. As you will see in Chapter 4, these prementalizing modes, while developmentally appropriate in preadolescent children, can be considered maladaptive "nonmentalizing" modes when they predominate mentalizing in adulthood. This chapter also emphasized that mentalizing capacity does not develop in a vacuum, but that the early caregiving environment (and later on, the social environment writ large) provides a critical laboratory for the practicing of mentalizing skills throughout development. Thus,

mentalizing is rooted in attachment relationships inside and outside the home environment. We showed that it is through parental mentalizing (or **reflective function**) and, in particular, marked mirroring, that the infant, child, and adolescent gain the capacity to mentalize. Parental mentalizing fosters in children a feeling of being understood, which in turn engenders epistemic trust in the environment as a source of social learning.

Guilford Publications
370 Seventh Avenue
New York, NY 10001
212-431-9800
800-365-7006
www.guilford.com