Mentalizing

Jon G. Allen, PhD

Abstract. Mentalizing, the process of making sense of mental states in oneself and other persons, plays a central role in psychopathology and psychotherapy. The author explicates the concept of mentalizing, highlights some factors critical to its development, and illustrates its clinical applications in the domains of trauma and depression.

I hope to persuade you that understanding “mentalizing” will sharpen your clinical thinking and practice. This will be a hard sell, because the general idea is not new. In some terms or other, you have been using the concept of mentalizing to think about your clinical work all along, not to mention the fact that you have been a full-fledged mentalizer since your early childhood. You are mentalizing whenever you are making sense of what goes on in the mind—your own mind and the mind of another person. Thus you mentalize continually when you conduct psychotherapy.

We mentalize so naturally that we usually take our mentalizing ability for granted. But our clinical practice sometimes reveals conspicuous failures in mentalizing. I worked with a young woman who had been sexually abused in childhood in the most degrading manner—merely used for gratification. This abuse was worse than humiliating, although it was certainly that; it was dehumanizing. In her young adulthood, stressful interactions often rekindled feelings associated with her earlier trauma. When she felt overwhelmed, she punched the wall, sometimes so violently she fractured her knuckles. In an early psychotherapy session, she told me that she had a powerful sensation in her chest and forearms and, along with the sensation, she had a strong urge to pound her fist into the wall. She had no idea what she felt emotionally, much less what had triggered the feelings, and even less what their origin was. Her experience was raw physical sensation coupled with an impulse to action. It made no sense.

Gradually, we were able to begin labeling some of the emotions the patient felt in conjunction with her physical sensations and impulses—agitation, fear, and anger. In one session when we were beginning to discuss the origins of these feelings, she casually and unconsciously put her hands around her throat as if she were going to strangle herself. This gesture was an unmentalized action, and I drew her attention to it. Plainly, the gesture could be interpreted as self-destructive. As we talked about its meaning, however, we associated the gesture with her inability to allow herself to have a voice. She could not speak out about the abuse; she had felt unheard persistently since then; and recent episode of self-injurious behavior was precipitated by a sense of having no voice in a significant decision in the family business. More metaphorically, her gesture of strangling herself suggested to me that she could not allow the feelings in her body to go up into her head, into her mind, to be mentalized. Gradually, we were able to bring some mental coherence to all this. She was able to identify the emotional feelings that triggered the impulse to punch the wall, and then she was able to relate the feelings to the meaning of some episodes of early abuse. In addition, she was able to understand the contemporary triggers of these feelings. When her feelings became meaningful, she was able to manage them more effectively, and her urge to injure herself began to subside.

As just introduced, mentalizing seems simple and commonplace. Commonplace it is, but simple it is not. This paper proceeds in three stages by (1) explicating the concept of mentalizing, (2) spelling out some of the conditions that foster and impede its development, and (3) illustrating its clinical applications to the domains of trauma and depression. This will be a thumbnail sketch; the ideas are elaborated elsewhere (Allen & Fonagy, 2002; Fonagy, Gergely, Jurist, & Target, 2002).

A word of warning: To get the concept of mentalizing off the ground, I will say a bit about the nature of mind, which plunges us into some heady philosophical concepts. The fact that it tangles us up in philosophy of mind makes mentalizing intriguing and vexing. For starters, mentalizing rests on the distinction between “persons” and “objects,” and I will begin there.
Then I will note briefly how two related technical concepts, intentionality and mental representation, are essential to understanding mentalizing. And we also have some technical ground to cover in developmental psychology, touching upon the ambitious project of fathoming how the mind—and mentalizing—comes into being in infancy. But my ultimate aim remains pragmatic, to show how highlighting mentalizing makes a difference in clinical practice. All this will become plain, I trust, in the section on clinical applications.

Conceptual Foundations
We mentalize when we treat others as persons rather than objects. Contemporary British philosopher, Peter Strawson, made this distinction with exceptional care and clarity, capturing the gist of mentalizing in the process: “We simply react to others as to other people. They may puzzle us at times; but that is part of so reacting” (1985, p. 21). These natural reactions, Strawson (1982, 1985) contended, are infused with moral attitudes and judgments; these reactive attitudes, in turn, are bound up with our experience of others and ourselves as free agents. Mentalizing in our inherently personal interactions with others, we are not usually objective, detached, or indifferent. Our natural interactions have a moral dimension to the degree that we approve or disapprove of others' actions; we feel grateful or resentful. These reactions are intrinsic to mentalizing interactively in all our relationships, those with our patients included.

Thus we mentalize, and we cannot help it:
Our general proneness to these attitudes and reactions is inextricably bound up with that involvement in personal and social interrelationships which begins with our lives, which develops and complicates itself in a great variety of ways throughout our lives and which is, one might say, a condition of our humanity. What we have, in our inescapable commitment to these attitudes and feelings, is a natural fact, something as deeply rooted in our natures as our existence as social beings. (Strawson, 1985, p. 33, emphasis mine)

Nonetheless, Strawson also argued that we can, temporarily and to some degree, detach ourselves from the mentalizing stance and thus from our natural reactive attitudes and judgments. We can view each other as objects, “as natural creatures whose behavior, whose actions and reactions, we may seek to understand, predict and perhaps control in just such a sense as that in which we may seek to understand, predict, and control the behavior of nonpersonal objects in nature” (1985, p. 34). Although Strawson contended that we do not naturally adopt this objective standpoint, he acknowledged that “perhaps we all edge a little way toward it from time to time” (1985, pp. 34-35).

Of course, in a non-mentalizing mode, we can treat others as objects in the most blatant ways. We can be exploitative, merely using others for our own ends. We can be callous and brutal in such exploitation, at the extreme, committing rape and murder. And, while we are prone to associating mentalizing with compassion, we must keep in mind that our mentalizing is by no means always benevolent. Psychopaths use their mentalizing capacities in the service of manipulation. Sadists mentalize most perniciously, with their keen awareness of inducing feelings of humiliation and terror.

Intentionality and mental representations
A definition: Mentalizing entails interpreting and responding to the behavior of oneself and others in terms of intentional mental states, such as desires, feelings, beliefs, and the like. Searle explained the meaning of “intentional” in this context:
The primary evolutionary role of the mind is to relate us in certain ways to the environment, and especially to other people. My subjective states relate me to the rest of the world, and the general name of that relationship is ‘intentionality’...the general term for all the various forms by which the mind can be directed at, or be about, or of, objects and states of affairs in the world. (1998, p. 85)
In short, intentional mental states are about something. Intentionality thus sharply divides mental beings from physical objects. A rock cannot be “about” anything; it just is. A mind, on the other hand, actively relates to something outside itself; a mind can think about a rock, another mind, or even itself.
The intentionality of mind—its “about-ness”—is bound up with its representational capacity. Mental representations, such as beliefs, represent something as being a certain way (Perner, 1991): I see her facial expression as being gleeful. Through this representational capacity, the mind assumes some autonomy in loosening its ties to reality. For example, our capacity to play is a harbinger of mentalizing (Leslie, 1987). We play with objects, and we play with ideas. To the child, a block of wood can be a truck. To the psychoanalyst, the whisky bottle can be a breast. Especially crucial to mentalizing is this ability to represent the same situation in multiple ways, to think of alternatives, and to adopt multiple perspectives. Mentalizing, we can imagine various ways in which others may think and feel, wondering why they do what they do, striving to make their actions intelligible. We also develop a capacity for meta-representation, being able to think about our own thoughts and feelings. Much of psychotherapy exploits this capacity for meta-representation as we engage our patients in reflecting on their actions and considering them from multiple perspectives.

Thus our mental flexibility is inherent in this representational capacity. We can be curious about possibilities, playing with reality, and playing with mental realities. In our best clinical work, we help our patients become curious about the workings of their mind and the minds of others. We engage them in the process of exploring various possible meanings of their feelings, dreams, and actions. When they lose this flexible curiosity, they are stuck, for example, in depression: “I failed. I’m no good. That’s the reality, the absolute truth. Period!”

Different aspects of mentalizing
These brief excursions into the nature of mind are essential to lay the groundwork for understanding mentalizing, but we cannot equate “mentalizing” with any and all mental activity. To reiterate, we mentalize persons, not objects. In theorizing about the structure of materials, engineers are highly engaged in mental activity, but they are not mentalizing. And when we treat persons as objects, we cease to mentalize. Nonetheless, the territory of mentalizing remains vast, and it is helpful to divide it up.

We will be hopelessly mired in ambiguity if we fail to keep in mind a fundamental distinction between mentalizing implicitly and mentalizing explicitly. Mentalizing explicitly, being explicit, is easiest to grasp. Mentalizing explicitly is a relatively conscious, verbal, deliberate, and reflective process—off-line, as it were. We are mentalizing explicitly when we think about what is going on in another person’s mind, for example, when we formulate clinical clarifications and interpretations. A simple example: “I wonder if you’re feeling put out with me right now?” Mentalizing implicitly, which is far more pervasive in our interactions, is a relatively nonconscious, nonverbal, procedural, and unreflective process, on-line as it were, akin to riding a bicycle. We are mentalizing implicitly when we empathize intuitively and nonverbally, “mirroring” others’ emotional states. We mentalize implicitly when we respond with a look of interest to what our patient just said, perhaps leaning forward a bit and raising our brow. We are barely aware of our implicit mentalizing much of the time, although we put a fair amount of effort into explicating the implicit, for example, drawing our patient’s attention to the meaning of his clenched fist. And we are mentalizing implicitly whenever we are emotionally engaged in interactions, with all the moral reactive attitudes Strawson emphasized.

Arguably, the distinction between implicit and explicit mentalizing is a matter of degree (Karmiloff-Smith, 1992), but this broad distinction is a place to start, and it is essential that we not equate mentalizing only with our conscious and reflective thought processes. But two further distinctions are in order. First, the object of our mentalizing varies: we mentalize the activity of ourselves, other persons, our relationships with other persons, and other persons’ relationships with each other. Second, we mentalize in relation to all manner of psychological processes: desires, feelings, beliefs, needs, motives, perceptions, attitudes, fantasies, dreams, hallucinations, delusions—the list is not quite endless, but almost.

The word, “mentalizing”
But why should we adopt the neologism, “mentalizing”? We already have in hand a slew of perfectly apt terms: empathy, insight, observing ego, subjectivity, transitional space, potential space, reflectiveness, mindfulness, and psychological mindedness. The term, mentalizing, joined the fray a few decades ago, when it was introduced into the psychoanalytic literature (Brown, 1977; Compton, 1983; De M’Uzan, 1973; Lecours & Bouchard, 1997). Morton and colleagues then significantly extended its application when they construed autism as a neurobiologically based failure of mentalizing (Frith, Morton, & Leslie, 1991; Morton, 1989). At
the same time, Fonagy and Target (Fonagy, 1991, 1995; Fonagy & Target, 1997; Target & Fonagy, 1996) further expanded its application by articulating how functional deficits in mentalizing, particularly those stemming from psychological trauma, illuminated severe character disturbance.

To perseverate, why “mentalizing”? One good reason is grammatical: we need a verb. We need a way to refer to a mental activity. Hence I prefer “mentalize” and “mentalizing” to “mentalization.” Alternatives like psychological mindedness, transitional spacing, or observing egoing simply will not do. “Empathizing,” although narrower in scope, captures significant portions of the territory of mentalizing, and contemporary theory and research on empathizing has much to teach us about mentalizing (Preston & de Waal, in press). From the grammatical perspective, a good alternative to mentalizing is Bogdan's (1997, 2000) usage of “interpreting” in the context of evolutionary biology. He defines the capacity for interpretation as “a competence that allows primates to make sense spontaneously and effectively of each other in terms of behavioral dispositions and psychological attributes, such as character traits, emotions, feelings, and attitudes” (1997, p. 1). Hence interpreting refers to the same domain as mentalizing. But interpreting has far broader connotations. Mentalizing, we interpret quirky actions, dreams, and Rorschach Inkblot Test responses. In addition to mentalizing, we interpret clouds, ancient texts, and tea leaves. Mentalizing is a restricted domain of interpreting: we just mentalize each other and ourselves (and perhaps our mammalian pets). Baron-Cohen’s (1995) term, “mindreading,” pinpoints the domain of social deficits in autism and also captures the territory of mentalizing but, absent the right context, mindreading falls prey to parapsychological connotations.

**Developmental Foundations**

We psychologists want more than philosophical distinctions; we want to know how mental capacities develop. The move from the “object” to “person” is acquired; it is not only a developmental achievement but also an evolutionary achievement. I advocated adopting “mentalizing” for a grammatical reason—we need a verb. This justification alone is hardly compelling. More substantively, the territory mentalizing carves out so neatly can be anchored in four domains of contemporary theory and research: evolutionary biology, neurobiology, theory of mind, and attachment. The concept of mentalizing has just the right scope for these connections, and its potentially strong ties to current research can put our venerable concepts of empathy, insight, psychological mindedness and the like on far more solid ground.

**Evolution and neurobiology**

Phylogeny is the broadest developmental perspective, yet one that is relatively peripheral to our clinical concerns. Nonetheless, the evolutionary perspective may help us appreciate the sheer complexity of mentalizing. Neither tool making nor foraging drove the evolution of our energy-consuming neocortex; rather, the evolutionary arms race was fought on the battleground of the mind-boggling complexity of our social reasoning (Bogdan, 1997; Byrne & Whiten, 1988; Humphrey, 1988; Jolly, 1988). Keeping track of the relationships among 150 persons in the typical human social network is no small feat (Dunbar, 1996). We need allies to survive and flourish. We must learn and monitor a multitude of relationships, all on shifting sands. Who is related to whom and how? Who are our friends, and who are our enemies? Far more complex still: who are each other’s friends and enemies? Researchers are delineating the “social brain” that evolved to cope with these challenges of daily social life (Brothers, 1997). This neurobiological endeavor has been bolstered by neuroimaging studies that are demonstrating specific patterns of cortical activity associated with mentalizing (Fletcher et al., 1995; Gallagher et al., 2000; Goel, Grafman, Sadato, & Hallett, 1995; Happe et al., 1996; Klin, Schultz, & Cohen, 2000).

**Developmental milestones**

Given an intact social brain, how does mentalizing develop? Much of the burgeoning developmental research in this area falls under the rubric of “theory of mind” (Carruthers & Smith, 1996; Gopnik & Meltzoff, 1997). Although the term, “theory,” seems a bit pretentious to characterize ways we learn to think about ourselves and each other in childhood, Gopnick (1996) pointed out that “it is not so much that children are little scientists as that scientists are big children” (p. 4). We are not behaviorists; we learn intuitively to postulate hidden mental causes to make sense of behavior, and behavior makes no sense otherwise. The physicist
fathoming subatomic particles started out as a psychologist fathoming the interpersonal world. The developmental research on theory of mind is vast, and I will merely highlight some key steps toward becoming a full-fledged mentalizer.

Our colleagues, George Gergely and his colleagues, unearthed a pivotal developmental process that hooks infants’ attention into the social world (Bahrick & Watson, 1985; Gergely & Watson, 1996, 1999; Watson, 1994). Initially, infants’ attention is captured by perfect response-stimulus contingencies, such as watching their limbs move. This attentional preference enables the infant to differentiate the self from the world, fostering a sense of bodily agency. In normal development, at about three months of age, a switch occurs, such that infants become more interested in high, but less than perfect, levels of response-stimulus contingency. They do something, and something related happens, just a bit later. Gergely (2001) aptly dubbed this newfound interest as a preference for stimuli that are “nearly, but clearly not, like me.” This clever gimmick of nature that switches infants’ attentional preferences has a profound consequence: infants become attentive to others’ responses to their actions, and thereby become ensconced in the social world. Autistic infants fail to make this switch, remaining fixated on perfect control over stimulation, failing to get hooked on social responsiveness (Gergely, 2001).

A perfect example of high but imperfect contingency is dear to the clinician’s heart: emotional mirroring. Indeed, Gergely and colleagues contend that the first step toward mentalizing entails mentalizing emotions (Fonagy et al., 2002; Gergely & Watson, 1996). Their focus on emotion at the inception of mind is consistent with Damasio’s (2003) view that “Feelings of pain or pleasure or some quality in between are the bedrock of our minds” (p. 3). Counter-intuitively, taking the Vygotskian perspective (Vygotsky, 1962; Wertsch, 1998), Gergely and colleagues propose that infants learn about their emotional states from the outside in. From our adult standpoint, it seems as if we know our emotions from the inside through introspection. For us adults, this may or may not be true—sometimes others may be more aware of our emotions that we are. But infants do not come equipped to diagnose their emotions on the basis of introspected feelings; they learn to make this connection via the emotional responsiveness of others.

In fact, what we glibly call emotional “mirroring” is more aptly construed as a process of social biofeedback (Gergely & Watson, 1996): the caregiver’s emotionally attuned responses to the infant’s states become a source of information to the infant about his or her internal states. This external feedback is, in effect, hooked up to the infant’s internal experience (e.g., perceptions of physiological arousal). Through such social biofeedback, infants obtain their first glimpse of what they feel. Thus begins a developmental process in which emotions become meaningful and ultimately can be re-represented in language.

Csibra, Gergely and colleagues (Csibra & Gergely, 1998; Csibra, Gergely, Biro, Koos, & Brockbank, 1999) also pinpointed another milestone on the infant’s way to full immersion in the world of intentional mental states, what Dennett (1987) famously characterized as the “intentional stance,” and what we call mentalizing. The precursor to the intentional stance is the teleological stance, which entails interpreting actions as efficiently achieving goals within the constraints of physical reality. Gergely and colleagues demonstrated that, by nine months of age, infants interpret behavior as rationally directed toward goals, even when the behaver is a computer-animated ball seen to be “jumping over” a barrier to make contact with another ball. The infant makes a quantum leap from the teleological to the intentional stance (see Gergely, this issue). This leap entails mentalizing the teleological stance, namely, interpreting rational, goal-directed behavior as guided by unobservable mental states, such as desires and beliefs. This is the leap from being able to predict behavior on the basis of observed regularities—as chimpanzees do and the behaviorists aspired to do—to being able to comprehend behavior in fully human terms.

Language acquisition is another arena that highlights the paramount developmental significance of mentalizing. Tomasello (1999) also took an interest in a nine-month turning point, the development of joint attention. He considered developing the capacity for joint attention as “the nine-month social-cognitive revolution” (p. 89) and described how joint attention ultimately fosters self-awareness:

As infants begin to follow into and direct the attention of others to outside entities at nine to twelve months of age, it happens on occasion that the other person whose attention an infant is monitoring focuses on the infant herself. The infant then monitors that person’s attention to her in a way that was not possible previously....From this point on the infant’s face-to-face
interactions with others...are radically transformed. She now knows she is interacting with another intentional agent who perceives her and intends things toward her. (p. 89) Communication requires joint attention, and acquiring language is one of the fruits of learning to relate to others as intentional beings: “Sounds become language for young children when and only when they understand that the adult is making that sound with the intention that they attend to something” (Tomasello, 1999, p. 101). Hence mentalizing implicitly—a sense of interacting with another mental agent—is the ground on which language develops.

**Attachment and arousal**

I have highlighted a few markers on the pathway to mentalizing just to hint at the complexity of the developmental process. Unfortunately, we cannot take for granted the developmental process of becoming a mentalizer, a person among persons. As already noted, mentalizing requires an intact social brain. But, as should be obvious by now, becoming a mentalizer also requires a nurturing relational context. Specifically, as recent research demonstrates (Fonagy, Redfern, & Charman, 1997; Meins, 1997; Meins, Fernyhough, Russell, & Clark-Carter, 1998), mentalizing develops optimally in the context of a secure attachment relationship. As Fonagy and colleagues (Fonagy et al., 2002) argue cogently, secure attachment is not only conducive to exploration of the outer world but also conducive to exploring the inner world—the mind of the self and the mind of the other. More precisely, and just as pertinent to the psychotherapy process as to early relationships with caregivers, secure attachment is conducive to exploring your mind in the mind of the other who has your mind in mind.

Just as secure attachment facilitates mentalizing, insecure attachment undermines it. Most glaringly, trauma in attachment relationships, such as abuse and neglect, confers a high risk for disorganized attachment (Lyons-Ruth & Jacobvitz, 1999; Main & Solomon, 1990; van IJzendoorn, Schuengel, & Bakermans-Kranenburg, 1999). In the context of trauma and attachment disorganization, mentalizing becomes frightening. Hence, as Fonagy and Target (1997) articulated, attachment trauma promotes a defensive withdrawal from the mental world. Being aware of the mind of the abuser confronts the child “with attitudes towards himself which are extremely painful to recognize: hatred, cruelty, indifference” (p. 693). Insofar as mentalizing—making sense of our experience—is among our core ways of regulating stress (Fonagy & Target, 2002), attachment trauma and mentalizing deficits escalate one another: the child avoids mentalizing, and the aversion to mentalizing undermines the child’s capacity to cope with the frightening relationship.

In part, secure attachment is conducive to mentalizing by virtue of facilitating an optimal level of physiological arousal (Field, 1985; Kraemer, 1999; Panksepp, Nelson, & Bekkedal, 1999). As already noted, mentalizing is among our most complex cognitive activities. Not surprisingly, mentalizing depends substantially on optimal prefrontal cortical functioning. The prefrontal cortex mediates “executive” functions, which include planning and temporal ordering of responses in the context of novelty and ambiguity (Goldberg, 2001). Nowhere do we encounter more novelty and ambiguity unfolding at a fast pace than in interpersonal interactions. Optimal prefrontal functioning, in turn, depends on optimal arousal. Arnsten (1998) neatly captured the gist of our vulnerability to being overloaded by arousal in her paper, “The biology of being frazzled.” As Arnsten and Mayes have delineated (Arnsten, 1998; Arnsten, Mathew, Ubriani, Taylor, & Li, 1999; Mayes, 2000), when arousal (mediated by norepinephrine and dopamine) exceeds a certain threshold, it is as if a neurochemical switch is thrown. This switch shifts us out of the executive mode of flexible responding into the fight-or-flight mode of habitual responding. The switch from contemplation to automatic action depends on a switch from predominantly prefrontal to predominantly posterior-limbic activation. Those who deliberated in the face of a leaping leopard did not live to pass on their genes. Those who ran, passed on a legacy: when we become highly aroused, our prefrontal cortex goes offline.

To put this simply, we all know the experience of being so anxious or frazzled that we cannot think straight. As an exceptionally complex cognitive process, mentalizing is quick to collapse in the face of excessive arousal. Moreover, persons with a long history of excessive arousal, such as would occur in traumatic attachment relationships, may become sensitized: their threshold for switching is lowered (Mayes, 2000). Such traumatized persons are more vulnerable to arousal, and their mentalizing capacities are more easily lost in the face of arousal. To mentalize, they need a secure attachment context conducive to optimal arousal.
In applying the concepts of attachment and arousal to clinical work, I find it helpful to draw a parallel between developmental competence and current performance. Developing competence in mentalizing requires an intact social brain, a secure-enough attachment history, and a frequent-enough experience of optimal arousal in relationships. Developmental competence is attained, to varying degrees, as a function of thoroughly intertwined psychosocial and neurobiological factors. Yet competence does not always yield adequate performance. A concert pianist can be paralyzed with stage fright. As it does in childhood, mentalizing in adulthood will best flourish in the context of a current secure attachment relationship and an optimal level of arousal. This facilitating context is just what we endeavor to provide in our clinical encounters.

Clinical Applications
We can best apply mentalizing to our clinical practice by shifting our focus from mental contents to mental processes, less concerned with what is on our patients’ mind than with the way they are using their mind. Focusing on mentalizing, I am now less worried about understanding the patient just right in my work as a psychotherapist, although I am not about to give up on that quest. Rather, I am more content just to be engaging the patient in the process of trying to understand. Thus I am not so much attempting to unearth any particular content (specific feelings, thoughts, or insights) as I am trying to create a capacity to mentalize more freely and flexibly—to foster involvement in personal relationships, as Strawson (1985) put it, and to foster greater awareness of mental states in self and others. I will illustrate the process of mentalizing in two clinical domains, trauma and depression. I will also note how the conditions that foster mentalizing in our patients apply equally to us clinicians. Finally, I will contrast the role of scientific psychology and folk psychology in our clinical practice, returning us to the opening discussion of objectifying our patients.

Treating trauma
The utility of the concept, mentalizing, crystallized for me in the context of treating trauma. Plainly, an interest in mentalizing is consistent with mainstream trauma treatment, for example, as Edna Foa (1997) articulated. Foa demonstrated the effectiveness of three essential ingredients in trauma treatment: emotional engagement (i.e., feeling the feelings), constructing a coherent narrative, and altering perniciously negative views of the self and the world, the latter including inadequacy (and sheer badness) of the self along with the dangerousness of other persons. Plainly, all this entails mentalizing as we understand it. But showing how one set of concepts (mentalizing) maps onto another (Foa’s account) does not show the advantage of adopting the concept of mentalizing. Although I might have come to the realization in other ways, it was only after Peter Fonagy got me enthralled (indeed, obsessed) with mentalizing that this idea crystallized in my mind: the goal of trauma treatment is not to put trauma out of mind but, on the contrary, to enable the patient to have the trauma in mind—to mentalize it, to make sense of it, and to bear it without succumbing to impulsive and self-destructive actions (e.g., drinking, bingeing, cutting, or trashing a room). This, of course, is the gist of “exposure” treatment, but the concept of mentalizing put this treatment in a new light (Allen, 2001); the procedure came to make more mental sense. Effective as it may be, to me “exposure” therapy—or its cousin “flooding” (Saigh, 1998)—has connotations of endurance: one learns to endure pain. No doubt, this idea is apt. Herman (1992), for example, usefully likened preparation for processing trauma to training for a marathon. But the concept of mentalizing shifted my emphasis from endurance and “desensitization” to mastery. No doubt, much of this work is grueling, especially for the patient, but also for the therapist. But it need not always be so. Once a patients have become able, in psychotherapy, to have the trauma in mind, I sometimes encourage them to keep traumatic images in mind, or even bring them to mind, rather than automatically putting them out of mind. Of course there is no point in dwelling on these images, except insofar as the patient learns not to be so frightened of them. Being able to have the traumatic experience in mind sets the stage for learning voluntarily to put it out of mind.

A young man who had been sexually abused was horrified when images of committing fellatio occasionally came to mind unbidden when he encountered men in his daily life. He became panicky and felt faint when he had these thoughts. We had worked in the psychotherapy to help him think, feel, and talk about the sexual abuse and then to calm himself by imagining
himself in a safe place in the woods which had been a refuge in his childhood. He began to feel in control of his mind in the therapy sessions. Then I suggested that he deliberately bring the dreaded images of fellatio to mind outside the therapy session for a brief time and then deliberately put them out of mind. He found that he could do so without undue distress, and he was no longer so vulnerable to being blindsided and panicked by these intrusive images. These are not new techniques, but my understanding of mentalizing shifted the way I thought about these techniques and my attitude toward them.

And we should not lose sight of the attachment context when we think about “exposure” and “desensitization” in trauma treatment. Enabling the patient to have the trauma in mind requires a secure attachment context and the optimal level of arousal (i.e., sense of safety and security) that such an attachment relationship fosters. Seeing the treatment in this way hardly changed my approach, but it gave me a more solid sense of conviction about the process, which has allowed me to do it with more facility. More easily, I encourage patients to think the unthinkable and to feel the unfeelable. These mental contents then become more thinkable and feelable, and then patients can achieve more voluntary control over their mind. Being able to have the trauma in mind, bearably, allows the patient to put it out of mind as well as to live in less fear of being blindsided by intolerable mental contents.

From this perspective on mentalizing, the various techniques recently developed for “processing” trauma (Allen, 2001) all rest on an ancient foundation—our basic humanity. I imagine the gist of trauma treatment to be as old as the origin of modern language (Corballis, 2002). Hence, a sketch of cutting-edge trauma treatment circa 50,000 years ago:

A: looks horror-stricken
B: asks, “What happened.”
A: tells the story
B: listens compassionately
A: mentalizes the trauma

B’s compassionate listening is therapeutic by virtue of A’s experience that B has her mind in mind (Fonagy et al., 2002). In Siegal’s (1999, p. 89) apt phrase, A “feels felt.”

Treating depression
A novel twist in current treatment of depression also illustrates the clinical application of mentalizing, albeit under the rubric of “mindfulness.” Teasdale and colleagues (Segal, Williams, & Teasdale, 2002; Teasdale, Segal, & Williams, 1995; Teasdale et al., 2000) have demonstrated that combining mindfulness meditation with cognitive-behavior therapy prevents relapse in patients with highly recurrent depression. Distinguishing two approaches to meditation helps elucidate the role of mentalizing in this approach (Goldstein & Kornfield, 1987). Concentration meditation entails focusing the mind, prototypically, on the process of breathing. Insight meditation, by contrast, entails letting the mind wander freely, adopting a non-judgmental attitude toward the kaleidoscopic flow of mental contents and states. Insight meditation also entails turning one’s attention toward rather than away from distressing thoughts and feelings, then observing natural changes in these distressing thoughts and feelings. Importantly, when observed non-judgmentally, distressing states change for the better; they wax and wane. Similarly, pleasant mental states change for the worse; they wax and wane.

The mindfulness meditation aspect of Teasdale and colleagues’ approach aims not so much to change the content of patients’ thoughts but rather to change patients’ experience of their mental states. As Teasdale (Teasdale et al., 1995) stated, patients are taught that “emotional disturbance is caused by thoughts and cognitions that are ‘mental events,’ not ‘realities’” (p. 38). Patients might think of “the mental state in which I view myself as utterly worthless” (p. 31). Segal and colleagues’ (Segal et al., 2002) thesis regarding the impact of their intervention captures the role of mentalizing, as we construe it: “Although the explicit emphasis in cognitive therapy is on changing thought content, we realized that it was equally possible that when successful, this treatment led implicitly to changes in patients’ relationships to their negative thoughts and feelings” (p. 38). Rather than becoming mired in depressive rumination (Nolen-Hoeksema, 1991, 2000), patients are able to let their depressing thoughts pass through, to construe them as mental processes rather than reflecting absolute truth. They are mentalizing,
aware of their mental states as such. Mentalizing, they can be “bummed out” temporarily without succumbing to depression. Thus they are demonstrably less prone to relapse.

In the same boat
Of course, we clinicians must mentalize to foster mentalizing in our patients. It is through our own mentalizing that we engage our patients in the process of mentalizing (and, conversely, through their mentalizing that they engage us in the process). We are in the same boat with our patients. We, too, must rely on an intact social brain, a secure attachment history, and an optimal level of arousal. We bring to the session our developmental competence, and our current state of mind (based on our feeling of security and level of arousal at the moment) may or may not be conducive to mentalizing performance. We, too, know the “biology of being frazzled” as our prefrontal cortical functioning goes off-line, giving way to our limbic propensities to fight, flight, or freeze responses. When our patients retreat from mentalizing, for example, by drawing us into reenacting traumatic attachment relationships, they provide us with direct experience of the biology of being frazzled. We too may feel threatened and unsafe. Plainly, the process of exploring the mental world requires a sense of safety and security for both parties, a sense of mutual trust. It takes two to mentalize therapeutically.

Reflections on folk psychology and scientific psychology
The billiards player relies on “folk physics,” and the mentalizer relies on “folk psychology”—in Bruner’s (1990) terms, “culture’s account of what makes human beings tick” (p. 13). Philosophers speculate that scientific psychology, informed by neurobiology, might someday render folk psychology obsolete (Lycan, 2003). For clinicians, this is a moot point. We psychotherapists, however well armed with scientific knowledge, must ultimately rely on folk psychology to make any use of it. We do best when we talk with our patients in plain language. And our implicit mentalizing is beyond language—and far removed from the scientific perspective.

I remember my first foray into conducting psychotherapy as a senior undergraduate student, with a few graduate courses under my belt, and headed for graduate school in clinical psychology. My advisor supervised me in conducting systematic desensitization (Wolpe, 1958) with a man who had a public speaking phobia. I had a scientific procedure: teaching relaxation, constructing a fear-stimulus hierarchy, and leading the client through the hierarchy while he maintained a state of relaxation—more relaxed than I, no doubt. I was a novice, but I could do that. After a few sessions, however, it became clear that my client just wanted to talk about his problems. I had no scientific procedure for that!

Off to graduate school, I held on to the illusion that there was a science to psychotherapy—some sort of algorithm. My mentors obviously knew it; I just did not yet “get it.” I never did get it, and the reasons are now clearer to me than ever. But I do not think I have been alone in my illusions, as the recent quest for manualized, evidence-based treatments suggests. We want scientific procedures for this daunting clinical practice: use your diagnostic decision tree, then prescribe the proper treatment, implementing your manual. Manual in hand (or better, in mind) you find yourself in the room with your patient. You also may be armed with a firm grasp of scientific psychology. How much good will the science do you? Imagine trying to interact with a patient (or anyone else, for that matter) on the basis of your knowledge of scientific psychology! Instantly you are thrown back on folk psychology—your ability to mentalize explicitly in ordinary language. And you will forge an attachment relationship largely on the basis of implicit engagement in mentalizing. Whether explicit knowledge of scientific psychology, psychoanalytic theory, or cognitive-behavioral therapy confers any advantage in the domain of what we traditionally call “the relationship” is an empirical question—perhaps a risky one to pursue. At best, with all our professional knowledge and skill, our mentalizing capacities are highly variable, as Diana Diamond and her colleagues (this issue) are discovering in their research on psychotherapy processes.

Yet, while emphasizing our dependence on our natural folk psychology and implicit mentalizing abilities, I am not suggesting that we toss aside all our professional knowledge. To be sure, professional knowledge (including treatment manuals) can be invaluable in providing a systematic therapeutic strategy, and extensive research attests to the effectiveness of evidence-based treatment approaches. Although it is a long way from informing me about what to do with the patient in the interactive moment, I could not imagine conducting psychotherapy without knowledge of psychopathology. And our treatments would quickly run
aground without a solid therapeutic frame (Gutheil & Gabbard, 1993), which is certainly a part of our professional heritage. But it is the work in the trenches that ultimately counts: the intuitive process of emotionally engaging the patient in mentalizing, using our own mentalizing capacities.

Relying too little on intuitive folk psychology and too much on scientific and professional knowledge always runs the risk of intellectualizing. Keeping too much science in mind runs the risk of objectifying the patient, the converse of mentalizing. I introduced mentalizing by contrasting persons with objects. And I defined mentalizing as interpreting the actions of oneself and others in terms of mental states. But this process of interpreting, as Strawson (1985) made plain, can be done in a detached manner:

To see human beings and human actions in this [detached, objective] light is to see them simply as objects and events in nature, natural objects and natural events, to be described, analyzed, and causally explained in terms in which moral evaluation has no place; in terms, roughly speaking, of an observational and theoretical vocabulary recognized in the natural and social sciences, including psychology. (p. 40)

Our patients are, of course, highly sensitive to excesses of this detached stance. Overly objectified, they will rightly complain of being analyzed, scrutinized, or put under a microscope. From this detached stance, we may formulate intellectualized “interpretations” that do the patient little good. Here we have the form but not the spirit of mentalizing. To be therapeutically effective, the mentalizing intentional stance must be grounded in emotional engagement.

Plainly, we must meld explicit reasoning and implicit mentalizing in effective therapeutic interactions. Scientific psychology, and professional knowledge more generally, will only be valuable to the extent that we can use this knowledge in the interactive moment to engage the patient in an attachment relationship that fosters a true meeting of minds. And there is no escaping it: this is moral engagement, in Strawson’s sense, with all its sentiments, judgments and “reactive attitudes”—resentment, gratitude, pain, suffering, satisfaction, frustration, and joy—sentiments that come naturally with any engaging human relationship. Without these sentiments, we could do no genuine “mirroring.” Detaching ourselves through explicit mentalizing (using our meta-representational capacities), we construe these sentiments as “countertransference,” which we aspire to use constructively to inform our work. But we do not and cannot remain detached. Weaving back and forth between objectivity and subjectivity, we are engaged in moral work, continually making emotional judgments (Nussbaum, 2001). Naturally, we must refrain from moralizing in the process (Grayling, 2002).

Of course, we must balance this inescapably natural mentalizing engagement with professional detachment, perhaps more than “edging away” from emotional involvement a fair amount of the time. Otherwise, we are liable to become embroiled in counterproductive countertransference. But edging away into scientific and professional objectivity, if it helps at all, will only help by informing and enriching our natural engagement. Psychoanalytic understanding, for example, can enhance mentalizing by directing our therapeutic attention to aspects of our humanity that we are prone to blocking from awareness. And we can use techniques of cognitive therapy to enhance our patients’ mentalizing by fostering more flexibility in their thinking.

Perhaps we should consider folk psychology and scientific psychology as being in continual interplay as we conduct our clinical work. Maybe this interplay captures the blend of art and science in psychotherapy, as we move back and forth between person-to-person emotional engagement and professional-scientific detachment. The latter is a relatively recent development in the therapeutic use of our humanity. If we developed the fundamental capacity to do this work at least 50,000 years ago, how much farther might we have progressed in the mere span of the past century?

References


