Journal of Personality Disorders, 19(6), 690–710, 2005

METAREPRESENTATIVE FUNCTIONS IN BORDERLINE PERSONALITY DISORDER

Antonio Semerari, MD, Antonino Carcione, MD, Giancarlo Dimaggio, MD, Giuseppe Nicolò, MD, Roberto Pedone, PhD, and Michele Procacci, MD

Many authors consider that patients suffering from Borderline Personality Disorder (BPD) are hampered in their ability to metarepresent, which is the correct ascribing of states of mind to oneself and to others and the reflecting thereon. Although the ability to mentalize is generally described as being uniform, various authors pinpoint problems which appear to be of a diverse psychological nature. Some describe difficulties in identifying emotions or a shortfall in their regulation, others identify a lack of integration between representations of self and those of others, and yet others focus on the failure to distinguish between fantasy and reality. In the present research all sessions during the first year of therapy of four patients suffering from BPD were tape-recorded and transcribed, and then analyzed using the Metacognition Assessment Scale (MAS), which is designed for the evaluation of the ability to metarepresent in clinical reports. The results support the hypothesis that there is a metarepresentation impairment in BPD but that it is more selective than was thought until now. In particular, such patients maintain their ability to identify internal states, whereas they are impaired in the integration of representations of self and others and in the differentiation between fantasy and reality.

A large number of observations have identified an impaired ability to reflect on mental states in personality disorders (PDs), involving difficulty in self-reflection (Ryle & Kerr, 2002; Westen & Shedler, 2000), accessing inner experience, recognizing others' mental states, and integrating different observations about own and others' behavior into coherent narratives. Some PDs are also described to be egocentric, unempathetic, and unable to attune to others (American Psychiatric Association, 1994; Shedler & Westen, 2004).

According to Livesley (2003), a difficulty in constructing integrated self-

From the Terzo Centro di Psicoterapia Cognitiva—Training School in Cognitive Psychotherapy, Scuola di Psicoterapia Cognitiva (SPC), Rome, Italy (A. S., A. C., G. D., G. N., M. P.); and Deptartment of Psychology, University of Naples II (R. P.).

Address correspondence to Antonio Semerari, Terzo Centro di Psicoterapia Cognitiva, Via Ravenna, 9/c, 00161 Rome, Italy; E-mail: terzocentro@iol.it

other representations is a core element in these pathologies. Most accounts of this impairment have focused on borderline personality disorder (BPD). Stimulated by Fonagy's ideas (1991), various authors have suggested that, to understand BPD, it is necessary to analyze the degree to which these patients are able to comprehend, describe, and process their own and others' states of mind (Dimaggio & Semerari, 2001, 2004; Liotti, 2002; Perris & Skagerlind, 1998; Ryle, 1997). However, descriptions of the inability to comprehend states of mind seem to refer to several distinct difficulties: what is usually considered a single function appears to be split into several sub-functions, each of which could be individually impaired. For example, difficulty in identifying and labeling emotions is thought to make subjects liable to states of emptiness, unable to regulate these same emotions, and therefore vulnerable to impulsive behavior (Fonagy, Gergely, Jurist, & Target, 2002; Linehan, 1993). Difficulty in understanding the processes involved in the shift between one state of mind and another is thought to involve a dissociation of these states (Ryle, 1997). Difficulty in maintaining a stable and consistent representation of others' and own mind is considered to foster chaos in relationships (Liotti, 2002). An inability to reflect upon one's own thoughts and affects may be caused by a difficulty in putting the narratives emerging in consciousness in order of relevance. This deficiency leads to confusion and an impairment of the ability to behave in a consistent manner (Dimaggio & Semerari, 2004; Fiore & Semerari, 2003; Semerari, 2001). A disorder in reflecting upon mental states may be also manifested in problems in differentiating between one's inner world and the world outside, and fantasy and reality (Fonagy & Target, 1996; Fonagy et al., 2002).

In summary, various mentalisation problems in reflecting upon mental states have, therefore, been ascribed to BPD patients, ranging from difficulties in monitoring and identifying emotions to an inability to integrate various states of mind and a failure to distinguish between one's inner world and external reality. Many of the clinical phenomena of BPD get attributed to these dysfunctions.

In this study we examined metarepresentation in BPD. Material from psychotherapy transcripts is used to explore whether there is a general impairment in the ability to comprehend states of mind in BPD; there are different dysfunction patterns; and these impairments are stable or fluctuate over time.

THE METAREPRESENTATION

The theoretical framework of our workconstructs on which our work is based comes from research into the metarepresentation functions carried out in the metacognition, theory of mind, and developmental psychology areas and from clinical literature about mentalization problems in BPD. *Metarepresentation* is the ability to form representations of representations (Sperber, 2000) and to use these representations to explain one's own and

others' visible overt behavior. Imagine the following scene: John, seated at a computer keyboard, gets up, takes a can of beer from the fridge, opens it and starts drinking it. An outside observer and John himself might explain these actions as follows: "John *wanted* a beer and *believed* there was one in the fridge; for this reason he went to get it." The want and the belief in John's mind constitute the first order representation, their representation in the observer's mind (or in John's when explaining his own actions) constitutes a second order representation, or metarepresentation (Falcone, Marraffa, & Carcione, 2003). There is much controversy about the origin of what is commonly called Theory of Mind (ToM). Some theories have an individualistic and nativistic nature and stress how they are the result of the progressive development of hard-wired structures in an infant brain, occurring at genetically predetermined stages, even if authors who support this position admit that their emergence is also affected by environmental factors (Baron-Cohen, 1995; German & Leslie, 2000). The alternative position is that the development of ToM is broadly linked to social interaction and happens gradually, with children not forming it on their own but for the most part learning it progressively in the course of their interactions with adults and peers (see Carpendale & Lewis, 2004, for an extensive discussion on the topic).

Whatever the relative importance of genetic or environmental factors, studies have suggested that an adult's normal ability to metarepresent comes from a gradual integration of several separate functions and skills. For Tomasello (1999), the first skill that emerges is the understanding of intentionality. At about 1 year old infants can grasp the teleological nature of actions and between 12 and 18 months they develop the ability to share intentions and goals and to engage in cooperative actions (Tomasello & Rakoczy, 2003). According to Baron-Cohen (1995), understanding intentionality develops from innate modular mechanisms; *module* means a sector of information processing, which is simple, innate, and specific to an informational domain (Fodor, 1983). Examples of modules that contribute to "mind reading" are the Eye Direction Detector (EDD) and the Intentionality Detector (ID), which reads movements in terms of goal directed actions (Baron-Cohen, 1995). These structures produce the input for the ToM Mechanism (ToMM; Leslie, 1994). The ID makes it possible to construct mentalistic propositions. In the example of John and the fridge, the process through which an observer translates behavior into a meaningful action is as follows: the representation *John looks at the fridge* is decoded by the EDD as *John directs his attention to the fridge.* ID gives an additional piece of information: *John is interested in the fridge.* Together, the two modular mechanisms provide input to ToMM for construing the hypothesis: *John desires something which he believes is in the fridge.* Fonagy and colleagues (2002) stress the role of social interaction as regards this position. In their opinion, interaction with adults in joint goal-directed activities is both the result of a child's innate inclination to understand and share intentionality and a condition for a further development of its meta-

cognition, which depends, therefore, to a great extent, on the quality of its relationships with caregivers.

Another skill that gets developed between 12 and 18 months is the distinction between fantasy and reality, indicated by a child's understanding of pretend play. In pretend play (Leslie, 1987) an object gets treated as if it was something else; for example, a banana as a telephone. According to Leslie, this involves the ability to distinguish between representations with a reference link to reality ("this is a banana") and representations that do not have this constraint ("this is a telephone"). This ability gets further developed over time and make it possible for children to distinguish between memories, dreams, fantasies, hypotheses, beliefs, and so on.

A crucial stage in the development of mentalization skills occurs around the age of 4, when a child learns to understand beliefs. The acquisition of this skill is indicated by the ability to solve the false belief task. In its original form (Wimmer & Perner, 1983) children are shown a short scene, in which an actor (Maxi) sees some chocolate put in a green box next to an empty blue one. While Maxi is absent, the chocolate gets moved to the blue box. The subjects then get asked: "Where does Maxi think the chocolate is?" (understanding of belief) and "Where will Maxi look for the chocolate?" (understanding of the beliefs guiding behavior). When Maxi comes back, the children are asked where he will look for the chocolate. To do the task correctly they need to understand that Maxi *believes* that the chocolate is still where he left it (in the green box); that is, they need to understand that Maxi has a false belief. The majority of research shows that almost all children under 3 years old give a wrong answer and claim that Maxi will look in the blue box, where it is in fact located. At around 3 years old a few children begin to answer correctly while at 4 it is the majority. Solving the false belief task is an indicator that a child possesses the basic skills necessary for becoming a good mindreader and is aware that behavior is driven by mental contents like beliefs and its ability to distinguish between the inner world and external reality. Autistic children typically fail in this task (Baron-Cohen, Leslie, & Frith, 1985).

Nichols and Stich (2001) maintain that the ability to understand others' intentions and beliefs (indicated, for example, by a correct answer in the false belief task) develops separately from the ability to recognize one's own intentions and beliefs, with knowledge of mental states in the first person deriving from a monitoring mechanism (MM) giving direct access to one's own inner states. They consider that this mechanism gets developed earlier than and separately from ToMM. For example, if 3-year-old children are asked to look at the contents of a box and are asked "Do you know what is inside?", they are able to solve the task without any difficulty, which is evidence that they are able to access their own thoughts successfully. On the other hand, access to others' thoughts is more limited: if they see another person looking at the contents of the box and the question becomes "Does he or she know what's in the box?", they then have significant difficulties (Nichols & Stich, 2001). Nichols and Stich also hypothe-

size that *detecting* ability, which is the identification specific mental contents in oneself, is distinct from the *reasoning* ability, in which one draws inferences about mental states and processes. In their opinion, one's reasoning skills are used in working out information about both one's own mental states and others'. Detecting, on the other hand, is part of the MM, used only in the first person.

To sum up, research into the development of metarepresentation suggests that it is made up of the following distinct skills: identifying intentions and desires, distinguishing between reality and fiction, recognizing beliefs, and detecting and reasoning. It is likely that future research will result in a further splitting up into separate functions. In this respect, metarepresentation seems to resemble memory, which is a global human skill (remembering), made up, however, of different forms and functions (working memory, procedural memory, autobiographical memory, semantic memory, and so on).

If metarepresentation consists of subfunctions which can be activated separately, it is possible to hypothesize that they can be impaired selectively in PDs (Semerari, Dimaggio, Nicolò, Procacci, & Carcione, 2005). For example, an individual might correctly identify his/her own thoughts (detecting not impaired) but not be able to reason about them (reasoning impaired). In this work we shall verify the hypothesis that only certain aspects of metarepresentation are impaired in BPD (Semerari et al. 2003).

METAREPRESENTATION DISORDERS IN BPD

Clinical observations of and research into metarepresentation functions in BPD are somewhat fragmented because some researchers treat metarepresentation as a single holistic function, whereas others focus on specific components. The most extensive evaluation of metarepresentation deficiencies in BPD was by Fonagy and colleagues (1996), who assessed the reflective function (RF) in a sample of nonpsychotic inpatients and suggested a low RF is not a risk factor on its own for BPD but predicts BPD when there has been abuse.

RF refers to the ability to understand mental states and perceive them in a coherent manner. The concept includes an awareness of one's own and others' states of mind, the effort to identify states of mind underlying behavior, the capacity to follow states of mind over time, and the ability to take account of other's state of mind. Fonagy and colleagues (1996) found that a lower RF was associated with BPD and histories of abuse; however, RF is a global concept so that a low RF indicates general metarepresentation problems, but does not tell us which specific functions are impaired.

An analysis of the literature on metarepresentative dysfunctions in BPD patients suggests that they have impairments in several of the subfunctions described in developmental psychology: identifying mental states, reasoning about them, and distinguishing between reality and fiction.

IMPAIRMENTS IN IDENTIFYING INNER MENTAL STATES

A dysfunction described in BPD is the inability to identify and describe one's own thoughts and emotions, with the most significant problems being in monitoring and identifying emotions. This impairment is equivalent to a failure to use the detecting function efficiently and, in general, to a problem in the MM (Nichols & Stich, 2001). Linehan (1993), hypothesized that this will arise in an invalidating environment, preventing patients from learning to label their negative emotional reactions accurately and leading to difficulties in communicating about their emotional states. An indicator of the ability to monitor inner states is alexithymia. Alexithymic individuals possess a limited vocabulary for describing their emotions and have difficulty in linking them to relational events (Taylor, Bagby, & Parker, 1997). Both of these difficulties involve aspects of what we have defined as metarepresentative skills. If this is the case, we should expect alexithymic traits and BDP to be associated. Research, however, has provided conflicting results. Bach, de Zwaan, Ackard, Nutzinger, and Mitchell (1994) analyzed the connection between diagnoses on Axis-I personality disorders (determined by Personality Diagnostic Questionnaire-Revised, PDQ-R; Hyler & Rieder, 1987) and alexithymia (assessed using the Toronto Alexithymia Scale, TAS; Bagby, Taylor, & Parker, 1994) in a sample of 182 outpatients. While they did not find any connection between Axis-I disorders and alexithymia, the latter turned out to have a strong link with schizotypal, dependent, and avoidant personality dimensions but not with the borderline dimension. In contrast, Berenbaum (1996) reported that abuse, borderline features, and a difficulty in identifying emotions were interconnected in a sample of 60 outpatients receiving psychotherapy.

IMPAIRMENTS IN INTEGRATING MENTAL STATES

Another deficiency ascribed to BPD is the inability to integrate mental contents and processes; i.e., to construct narratives putting together various aspects of one's subjective experience. In integrated narratives individuals, as well as storing various scenes from their autobiography in working memory, display a good ability to reason (Nichols & Stich, 2001) and to make plausible explanations for transitions between states; for example, "I was calm in the morning, but then, because of how my boss criticized me, I became irritated and this led me to reply snappily to my friends." The concept of *non-integration* gets used in clinical practice, however, to indicate processes involving different phenomena. A first category of definitions refers to non-integration *between* different states of mind. By this we mean that there are states of mind that are internally coherent but in conflict with or contradictory to other states of mind. They alternate with each other, without the individual managing to reflect on the conflict between them so as to integrate them. Kernberg (1975) ascribes this phenomenon to primitive splitting and projective identification. Given that the

splitting involves the positive and negative aspects of the object and of self, this concept stresses the fact that the non-integrated states are opposites, with representations of self and other that are all positive or all negative; however, non-integration does not necessarily involve opposite states. Individuals may, for example, swing between states in which they feel vulnerable victims and others in which they portray themselves as persecutors and abusers, without putting them together in an integrated way. Ryle (1997), for example, suggested that in BPD childhood experiences of maltreatment or abuse impair the ability to understand the processes governing the shift between one state of mind and another. According to Ryle, each of us has a stock of role representations regulating and modulating our relational behavior. A role representation gets accompanied by recurring emotions and thought themes. Patients with BPD do not develop conscious thoughts about their various role representations and the shifts between them. Consequently, they experience states of mind that are dissociated and segregated from each other and they tend to swing or shift from one to another in a rapid, chaotic, and poorly modulated way.

A second definition of non-integration focuses on disorganization in one single state of mind. Several authors have linked this tendency to the formation of a disorganized internal working model (Fonagy & Target, 1996; Liotti, 1999). When faced with a frightened and/or frightening caregiver, a child tends to introject the former's sensations of fear and anger into its self-structure, together with a self-image which is, at the same time, frightened and frightening. Such experiences make it likely that these individuals, when adults, develop multiple and contradictory representations of other's mind, in which they simultaneously ascribe a persecutory attitude, a willingness to help, and a need for help and protection, together with a complementary multiple and contradictory representation of self (Liotti, 2002).

A further way of defining the non-integration of single states of mind is the inability to put thoughts and emotions in order, leading to more chaotic experience. From this point of view, difficulties with reflecting on the contents of one's mind can lead to an uncontrolled throng of different thought themes and feelings, all intense and without a structure or a hierarchical order based on their meaning and importance. This results in a disorganized narrative (Dimaggio & Semerari, 2001; 2004).

Non-integration, therefore, is used with different meanings and may refer to different phenomena. In order to discover a common thread we need to understand what function integration performs. Human beings' thoughts and affects are intrinsically complex and contradictory. An ability to reflect on them is, therefore, needed to put them in order of importance and priority. This organization in turn provides direction and continuity to actions. Integration deficiencies tend to lead to a lack of consistency in behavior. All the types of integration deficiency we have described involve a loss of mental order, making actions contradictory, chaotic, or lacking in direction.

IMPAIRMENTS IN DIFFERENTIATING BETWEEN REPRESENTATION AND REALITY

A third dysfunction described in BPD is the inability to distinguish between representations based on fantasy and those based on reality; i.e., to use pretend play correctly (Leslie, 1987). Severely disturbed attachment relationships can impair this skill and this is a risk factor in borderlines. Fonagy and Target (1996) found that a young child tends to swing between two thinking modes: the *equivalent* mode and the *pretend* mode. In the equivalent mode mental events are equivalent to events in the outside world, in terms of emotional force, causality, and implications. In this case, a child hitting a rag doll with a stick might cry its heart out after damaging it, as the child feels to blame for having injured it. In the pretend mode, on the other hand, ideas and feelings get placed in the "as if" category and are perceived as being different from reality, so that damage to the rag doll leads the child to stop playing with it, but without any feelings of sorrow or guilt for the injury "inflicted." It is only with the help of another mind that a child learns "to play with reality," that is to consider its inner reality to be, at the same time, distinct from and linked to the outside world. A child's experience with parents who are frightened by its emotional manifestations, or mistreat or abuse it, is considered to damage this process. As a result the child cannot learn to reflect on his/her state of mind, and when the child becomes an adults ther will be the inclination to experience his/her own mental states as unadulterated registrations of external reality.

At least three types of metarepresentative disorder, therefore, have been described in BPD: monitoring, reasoning and integration, and differentiation. If this is the case, one can hypothesize either that there is one single metarepresentation disorder, of which the various authors have described different facets, or else that the clinical descriptions refer to different disorders. Our hypothesis is that metarepresentation derives from an interaction between different subfunctions and that these can be impaired separately from each other. If this is true, there can, in pathological situations, be two types of uncoupling between functions: a same patient with some functions working normally and others impaired, for example with monitoring functioning but not integration; and different patients, or groups of patients with different diagnoses, differing from each other as to the type of functions impaired, with, for example, some having an impairment of monitoring but not integration and others vice-versa.

METHOD

We took the transcripts of all the first year of psychotherapy sessions of four patients diagnosed primarily for BPD using the SCID-II for DSM-IV (First, Gibbon, Spitzer, Williams, & Benjamin, 1997). This gave a total of 138 sessions. Metarepresentation was evaluated using the Metacognition

Assessment Scale (MAS; Semerari et al., 2003). In addition, Axis I was evaluated by a therapist on the basis of DSM-IV criteria and confirmed by two expert clinicians.

PARTICIPANTS

The four individuals were all women aged between 24 and 38 ($\mu = 29$) and we will refer to them with the abbreviations Aus, Mag, Lor, and Ser. Three of them had a high school leaving certificate, and one (Mag) a conservatory diploma. When they started therapy none of them had a stable sentimental relationship. Lor and Ser did odd jobs, Mag had a permanent job, and Aus was unemployed and lived with her parents. One (Aus) also met the criteria for paranoid personality disorder, and another (Lor) met the criteria for avoidant personality disorder. On Axis I three patients were suffering from a major depressive episode (Mag, Aus, Lor), and the fourth (Ser) from a not otherwise specified dissociative disorder. This last patient reported, moreover, a history of recurrent abuse during childhood. During the period under consideration only one of them terminated the therapy with a successful outcome, as calculated using BPRS (Nicolaou et al., 1995), AMDP (Guy & Ban, 1982), and GAS (Endicott, Spitzer, Fleiss, & Cohen, 1976), while the others were still under treatment at the end of the period.

METACOGNITION ASSESSMENT SCALE

The MAS measures variations in individuals' metarepresentation functions. MAS features three sections referring to how much individuals are able to, respectively, think of their own mental states (Understanding of One's Own Mind—UM) and of others' (Understanding of Others' Mind—UOM), and to master psychological problems (Mastery—M). To support the scale's construct validity the metacognitive functions were investigated in literature on mentalization and attachment theory (Main 1991; Fonagy, Moran, Steele, Steele, & Higgitt, 1991) and metacognition (Flavell, 1979) for UM, on Theory of Mind (Baron-Cohen et al., 1985; Premack & Woodruf, 1978) and metarepresentation (Frith, 1992; Sperber, 2000) for UOM, and on metacognitive regulation (Nelson & Narens, 1990) and mastery of problematic experiences (Stiles, Meshot, Anderson, & Sloan, 1992) for M. In this study we use only the UM part of the scale, which measures how much an individual is capable of recognizing, distinguishing between, integrating, and reasoning about his/her own mental states, as measured by a set of subfunctions: monitoring, integration, and differentiation.

Inter-rater reliability was calculated using Kendall's W coefficient. Agreement was calculated for all sections with submission of two sample sessions from two patients to three independent judges and we obtained $W = .935$ for the first sample patients and $W = .931$ for the second sample patients ($p < .01$). As regards concurrent validity, the MAS was correlated (Lysaker et al., 2005) with data obtained by the Scale to Assess Unaware-

ness of Mental Illness (SUMD; Amador & Strauss, 1990), which measures the insight of individuals as to a specific aspect of their existence (awareness of disease). We chose this instrument because, to be aware of their own mental disorder, individuals need to have access to their inner states. In a sample of schizophrenics the results showed a positive and significant correlation between the two scales ($r = .27$, $p < .05$), as expected. Other evidence supporting MAS concurrent validity can be found in Glick, Salvi, Stiles, and Greenberg (2004) and Carcione and colleagues (2004), where 12 psychotherapy sessions with a patient suffering from major depression were analyzed by the MAS and the Assimilation of Problematic Experience Scale (APES; Stiles et al., 1992). Both scales measure the degree of access to inner experience. In the sessions where patients showed an increase in the capacity to assimilate problematic experience (measured by APES) their metacognition (measured by MAS) improved too.

ASSESSMENT AND RATING

In order to explain how the assessment of metacognitive functions is made, we present below some examples from psychotherapy sessions which represent successes and failures of the functions examined in the current paper. A rater has to identify, in each unit (each piece of speech by a patient between two interruptions by a therapist) a patient's attempt to exercise metarepresentation and, if this is the case, whether it has been used correctly and congruously or not: in the first instance we can talk of *success* (marked with a *yes*) and in the second of *failure* (marked with a *no*). Consequently successes and failures are evaluated separately for each function.

Monitoring[1] consists of the ability to identify the thoughts and emotions making up one's inner states and to grasp the emotional significance of the signals being sent out by another. Monitoring skills also include the ability to grasp the immediate effect that events have on emotional states and the relationships between thoughts and emotions. This function has two aspects: (1) identification, being the ability to perceive the emotions or ideas making up a state of mind (for example, "I felt sad"; "I imagined I was on my own in a tumbledown house"); and (2) relating variables, being the ability to grasp the relationships between the various components in a state of mind (for example, "I was sad because I was thinking about how I'd failed") and the causal relationships between external events and inner states (for example, "Linda's coldness made me think that she didn't care about me"). Consider the following passage:

1. Our definition of metacognitive monitoring includes certain basic activities that are not taken into account in other scales. For instance, the Reflective Scale considers the statement "I was angry" to be a nonreflective one, whereas the MAS interprets it as an identification operation.

P: To say to always go with a coat for six months, the same trousers for months, even though I say "no, it is time for me for changes," then after it is always the same story . . . when the moment arrives. This is also true when I told you "I want to change my direction to come here, I want to propose another thing." But then here arrives that moment . . . always that thing. It is as if there was an invisible force that carries you always on that same track.

In this case a failure in monitoring (*No*) is marked because the patient does not describe the thoughts or emotions guiding her behavior and it is therefore impossible to grasp why she always wears the same clothes or always travels in the same direction and sits in the same place.

Integration is defined as the ability to reflect on different mental states and/or contents giving them an order and hierarchical relevance. It consists of a dialogue inside ourselves in narrative form and gives a sense of continuity to the private and interpersonal aspects of our self (Bruner, 1986; Dimaggio & Semerari, 2001; Hermans & Kempen, 1993; Neimeyer, 2000; Sarbin, 1986). Following is an example of a success:

P: That's just like me. I start out with great enthusiasm and I'm sure I'm going to manage. Then the effort involved makes me less sure of success. I find the thing falling to pieces before me and I lose interest. At this point I depend wholly on others' giving me their approval if I am to get going again.

What follows is an example of non-integration:

P: I couldn't get to sleep. I was thinking about trying to pluck up courage and of how I don't manage to. I just wanted to sink into the arms of someone willing to take me. I wanted to wake up my parents and spend a little time with them or else go to the sitting room. They'd have heard me and called me. . . I imagined that they'd have got angry, because I'd waken them up, and have rebuked me and I didn't know how I'd have reacted.

T: So you thought they'd have criticized you?

P: Yes, I imagined that they'd have said, "Even during the night you have to be a nuisance!" and it annoys me. I feel the anger inside me because it's not fair. I'd like things to go differently and instead I feel guilty because I make them suffer.

The patient has no difficulty in describing her emotions and thoughts, but the overall representation of her own and others' minds appears contradictory. Her parents are seen at the same time as welcoming, suffering, scornful, and critical. The patient's emotions appear unsteady and her behavior lacks coherence.

Differentiation is the ability to recognize that the contents of representations are subjective events and merely hypotheses, and not objective descriptions of reality, and that they do not have a direct influence on it. A difficulty in looking critically at one's beliefs and in considering them hypothetical representations and/or a belief that thoughts may directly influence reality signals a malfunctioning. For example, here is a BPD patient after ripping up his girlfriend's clothes during a fit of jealousy:

P: I couldn't remember where they'd been sitting the evening before. Now I know that they were at opposite ends of the table (he indicates where), him here and her there. But that morning I couldn't remember . . . I could see them close to each other and chatting, while I was left out."

During the session the patient recalls the evening differently to how he imagined it in his fit of jealousy. In the session he is able to distinguish between his fantasy (seeing them sitting close to each other at a table) and reality (sitting at opposite ends). During his fit of jealousy the patient treated his fancy about being left out and deceived as a real memory and this led to an acting-out.

RESULTS

Figures 1 to 3 show the percentages of failures and successes in the use of monitoring, integration, and differentiation functions, respectively. The light columns stand for the percentage of successes (computed as: *yes/yes+no*) and the darker ones indicate failures (computed as: *no/yes+no*). Moreover, a linear trend line shows a function's course during the observed period.

MONITORING

Figure 1 shows the trend in monitoring function: Mag and Aus have practically no failures. In Lor they are to be found in only the very earliest sessions, but are significantly less than her successes. A longitudinal analysis shows a progressive decrease in failures until they almost disappear in the last sessions. On the other hand, in Ser's case, where there is a dissociative disorder and a history of abuse, we find a more substantial level of failures—albeit less than the successes—throughout the period analyzed and with a tendency to increase as therapy proceeds.

INTEGRATION

The trends in integration are different. As we can see from Figure 2, all four patients share a clear impairment in this function, albeit with variations in the trend over time. In Mag, Aus, and Lor there is a fluctuating trend: the earliest sessions contain mainly failures, followed by other sessions with mainly successes. This alternating trend gets repeated throughout the period considered. In Mag, the only patient to complete her therapy in the period in question, and, to a lesser extent, in Lor, there appears to be a stable trend toward improvement in the second half. Aus still has a high number of failures, exceeding successes, at the end of the period. In Ser, on the other hand, there appears to be a more constant and more severe integration deficit, and the failures are higher than successes in almost every of the session.

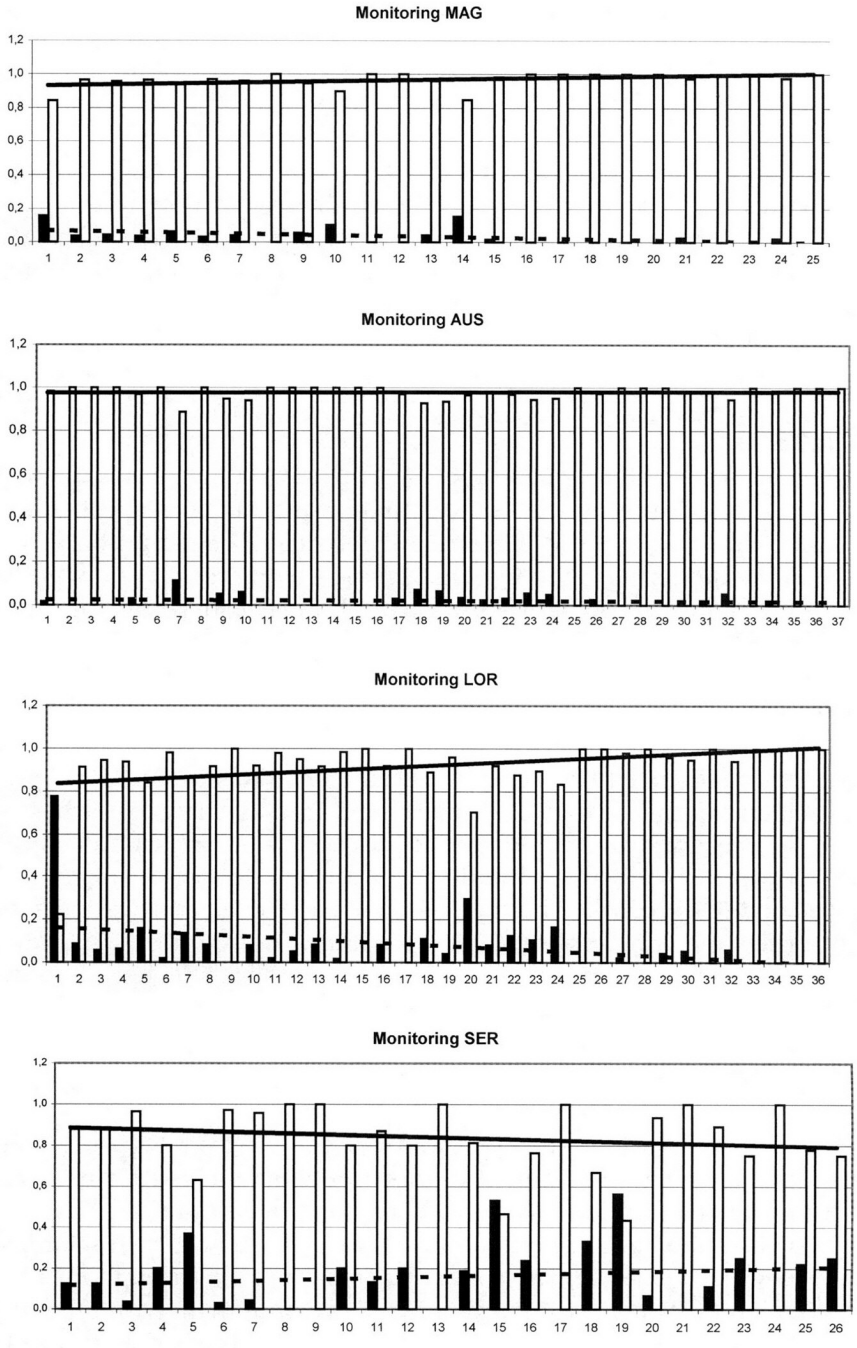

FIGURE 1. Failures and successes in monitoring functions during the 1st year of psychotherapy.

Note. The *x axis* indicates the sessions' number. On the *y axis* the white columns indicate the rate of successes (*yes*) and the black columns indicate the rate of failures (*no*) in monitoring internal states pointed out in each session. The solid and the dotted linear regression lines indicate the trend of success and failure, respectively.

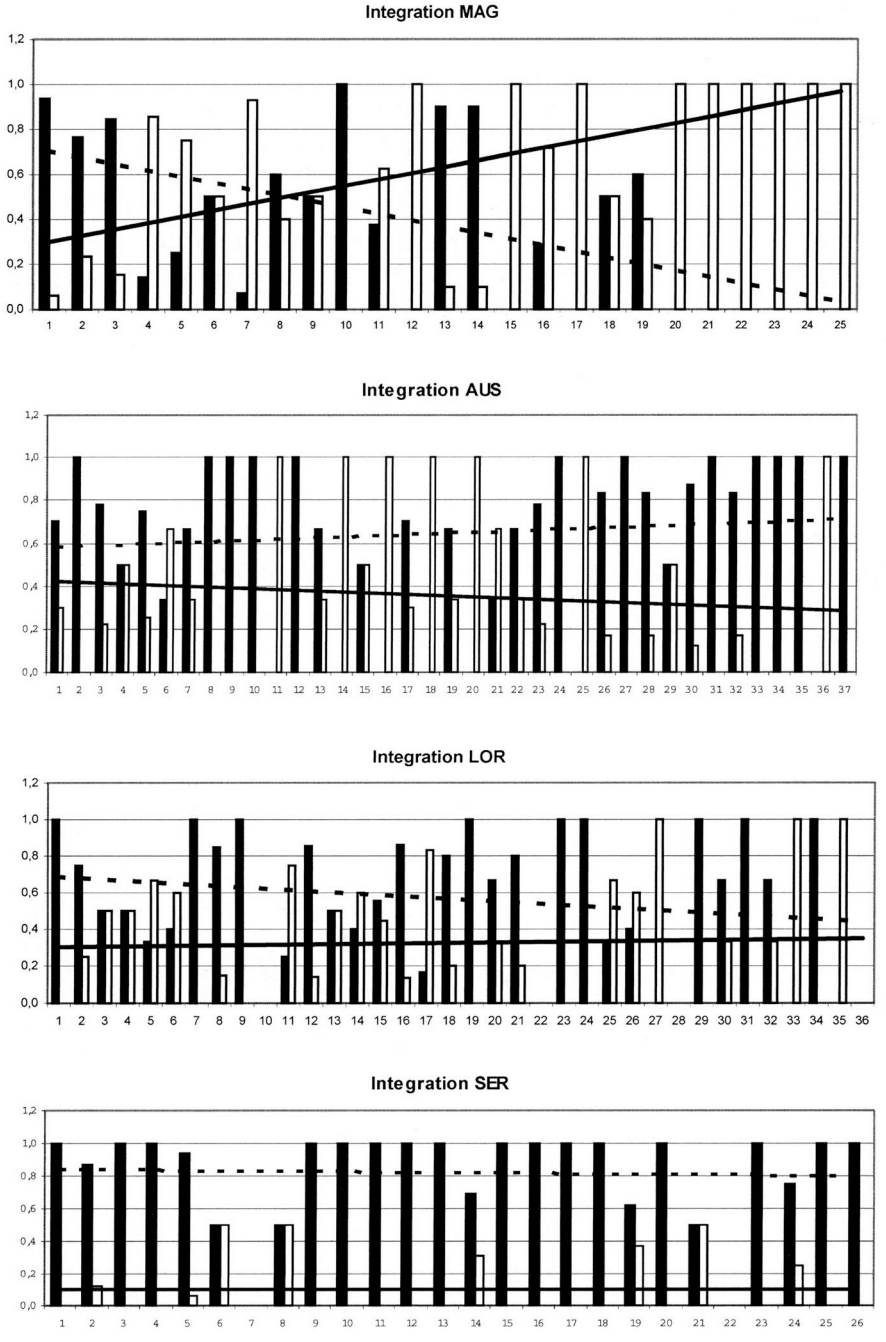

FIGURE 2. Failures and successes in integration function during the 1st year of psychotherapy.

Note. The *x axis* indicates the sessions' number. On the *y axis* the white columns indicate the rate of successes (*yes*) and the black columns indicate the rate of failures (*no*) in integrating internal states pointed out in each session. The solid and the dotted linear regression lines indicate the trend of success and failure, respectively.

DIFFERENTIATION

All the patients examined displayed an impairment in this function too (Figure 3). In three cases, Mag, Aus, and Lor, there is a fluctuating trend, similar to the one shown for the integration function, with a swing between sessions, or groups of sessions, with a majority of failures, followed by other sessions with mainly successes. In Ser it is striking how little the use of this function is noted; there are even some sessions in which neither successes nor failures can be observed. We can hypothesize that the level of her integration impairment, which is decidedly more severe than in the other patients, makes her conversation confused and chaotic to such an extent that it is difficult to evaluate her ability to distinguish between her imagination and reality. It has to be noted, however, that the number of failures tends to decrease as therapy proceeds.

DISCUSSION

As the analysis covers a long period, it is possible to pinpoint that the first three patients (Mag, Aus, and Lor) have similar profiles as regards type of impairment and trend in metarepresentative functions. They all keep the ability to monitor: that is, they are able to tell their therapist about their thoughts and emotions and also to describe the immediate causes of these thoughts and emotions. Vice-versa, we can see how all three clearly have problems in providing an integrated description of their mental scenery (an impaired integration function) and tend to consider their own representations as being a true reflection of reality (an impaired differentiation function). As a result, their representations of their states of mind are confused and lacking in critical detachment, although they are not opaque. Ser has an integration deficit too, although in her case it is more substantial and with a more constant trend over time. She also has problems in differentiating, even if this is less easy to note. She is different from the others because she is the only one with significant monitoring problems. To sum up, consistent with Kernberg (1975; Clarkin, Yeomans & Kernberg, 1999), Ryle (1997), Liotti (2002), and Dimaggio and Semerari (2004), all four patients suffer from a severe integration disorder. We would recall that this last result has been obtained using a definition of integration wide enough to include all the various definitions by these authors. It has to be noted, however, that this valuation does not take into account the current debate about the causes of these problems; for example, if it is due to a defensive mechanism (as proposed in Kernberg's theory) or to a functional deficit (as claimed in Liotti's and Ryle's theories).

The explanation that all four individuals have problems differentiating is consistent with the hypothesis of Fonagy and Target (1996) that borderline patients have a tendency to lose their ability to distinguish between fantasy and reality.

The fact that three out of the four patients have no problems in identify-

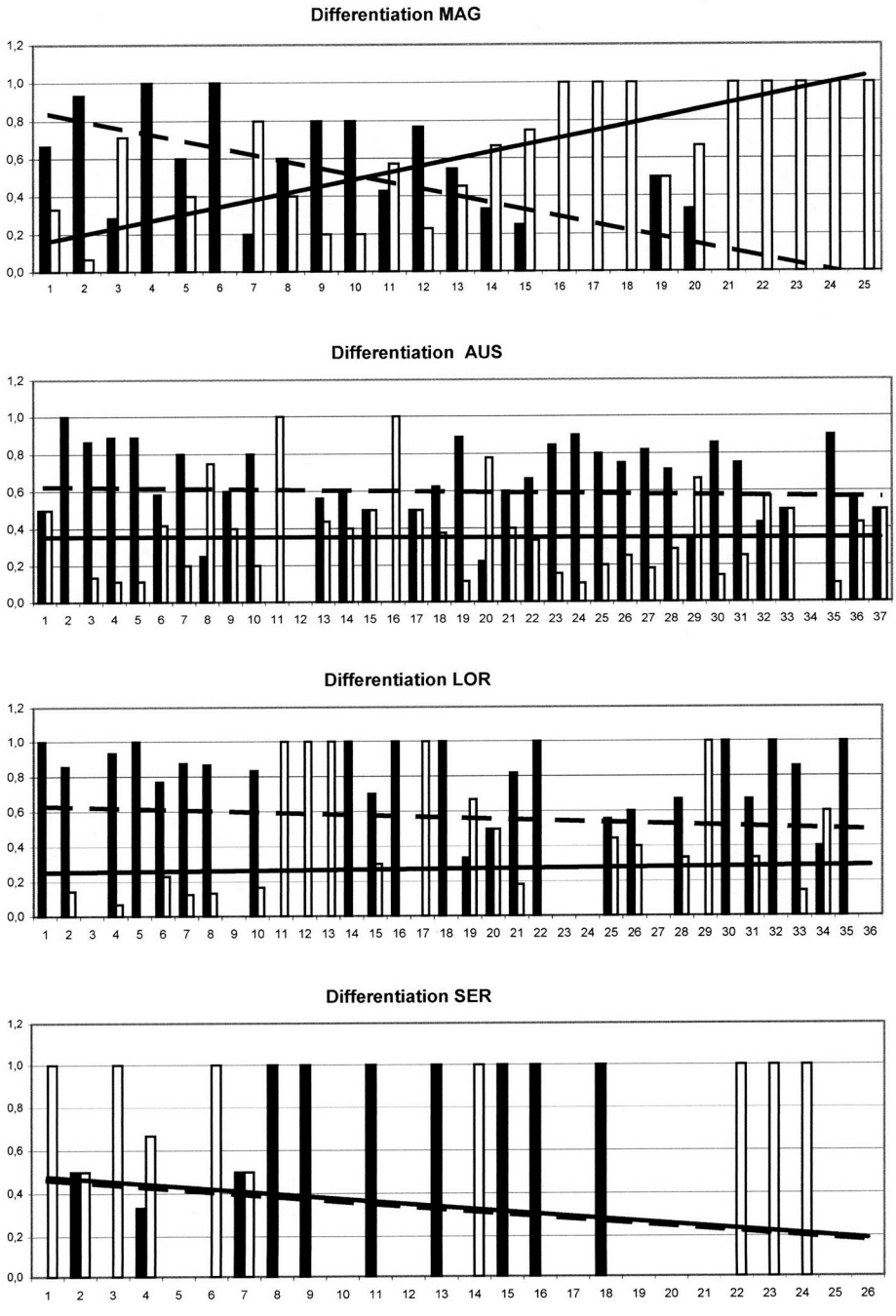

FIGURE 3. Failures and successes in differerntiation function during the 1st year of psychotherapy.

Note. The *x axis* indicates the sessions' number. On the *y axis* the white columns indicate the rate of successes (*yes*) and the black columns indicate the rate of failures (*no*) in differentiating internal and external reality pointed out in each session. The solid and the dotted linear regression lines indicate the trend of success and failure, respectively.

ing the emotions and ideas composing their inner states does not support the theory that a monitoring impairment is a typical BPD dysfunction, and the emotional disregulation described by Linehan (1993) does not, therefore, derive from difficulties in identifying and labeling emotions correctly. On the contrary, it is possible that the intensity with which borderlines experience affects contributes to a reinforcement of monitoring skills. The fact that monitoring is not impaired in BPD patients but they nevertheless suffer from highly impaired integration and differentiation supports the hypothesis that there is not an overall metarepresentation deficiency in severe personality disorders, in contrast to previous hypotheses (Fonagy et al., 2002). There may be malfunctioning in certain subfunctions, while others are, instead, undamaged (Dimaggio, Semerari, Carcione, Nicolò & Procacci, in press; Semerari et al., 2003, 2005). Ser's case, with its monitoring malfunctioning, is in line with other findings (Berenbaum, 1996; Linehan, 1993; Searles, 1988) that *some* borderline individuals have difficulty in identifying their inner states. This suggests that there can be different malfunctioning profiles within a single diagnostic category and confirms that it is important to identify each patient's specific deficiencies. What we are saying is in line with the idea that the best definition of personality disorders is in the form of descriptive prototypes (Westen & Shedler, 2000). The data emerging from our research suggest that the BPD prototype includes integration and differentiation dysfunctions, but that certain patients may also display, as a less core trait, a monitoring impairment.

Widening the sample in this research would make it possible to evaluate whether a monitoring impairment is part, even if marginally, of the prototype or is an idiographic element. An alternative hypothesis is that, given the high comorbidity of other PDs with BPD (as, for example, in the sample considered here), it is likely that differences in metarepresentation disorders are related to the variations in diagnoses and codiagnoses. If this is so, the greater severity of Ser's condition probably reflects the fact that she also suffers from dissociative disorder and has a history of abuse. This would be consistent with the evidence indicating that a history of abuse is one of the factors impairing metrepresentative functions the most (Berenbaum, 1996; Fonagy et al., 1996) and that there is a tie between alexithymia and dissociative disorder (Modestin, Lötscher, & Erni, 2002).

Research involving the therapeutic process has the advantage of giving us data about the trend in variables over time. In this example, with the exception of Ser, who has a stable integration dysfunction, there is a significant alternation in the other patients between moments where they have severe problems and others where they function fairly well. This fluctuation in functioning indicates that we are not witnessing a deficit that is unchanging and stable over time, but a form of malfunctioning subject to sudden alterations and dramatic improvements. It is likely that life events and the relational contexts in which these patients live have a strong impact on their skills: positive events, either in daily life or in the therapeutic relationship, can lead to improvements, which are then succeeded by se-

vere deteriorations caused by negative events. Validation (Linehan, 1993) or experience-sharing (Dimaggio & Semerari, 2004) interventions might, therefore, lead to improvements like those we have seen. A preliminary qualitative analysis of some PD patients' session transcripts shows that interventions in which therapists stress that they share some aspects of their patients' experience are often followed by rapid improvements in those patients' metarepresentation (Fiore, Pedone, Petrilli, Popolo, Carcione, & Nicolò, 2004).

The results seem to support our hypotheses that (a) metarepresentative skills are impaired in BPD; and (b) that the impairment does not involve metarepresentation as a whole, but hits only certain aspects of it selectively. The main impairments are in the abilities to integrate aspects of experience and to differentiate between fantasy and reality. In three of the four patients, on the other hand, monitoring was undamaged; (c) the impaired functions tend to fluctuate between periods of malfunctioning and others in which they function quite well. Metarepresentation deficiencies are, therefore, not stable in BPD and probably depend on trends in interpersonal relationships. In particular, moments of stress may cause a deterioration, with an example being severe dissociative episodes in which the fantasy/reality distinction is completely lacking, even if only temporarily.

We consider that this type of research has a number of advantages. A qualitative analysis of the psychotherapeutic process provides information not obtainable through a self-report or semistructured interviews. Compared with self-reports, in transcript analysis the material is elicited by an expert clinician rather than by an interviewer, often one with no clinical skills. Moreover, in self-reports patients are required to exercise those very mental functions that are impaired (Shedler & Westen, 2004).

Furthermore, an ongoing marking of functions in session transcripts provides information on their trend over time, demonstrating, for example, how a metarepresentative malfunctioning may not be a fixed and stable deficiency but subject to variations and fluctuations. When identifying trends over time, it is also possible to pinpoint, in the same transcript, the environmental, therapeutic, and relational factors influencing them. To obtain similar information with semistructured interviews, it would be necessary to repeat them for such a number of times that the advantages in terms of time would be lost or the answers to the interviews would probably become stereotyped.

The present study has a number of limitations. Firstly, this kind of research is very time-consuming, compared to self-reports, which are relatively easy and fast to administer. Secondly, the sample in our study consists of four women, treated by therapists from the same school and with similar experience. It needs, therefore, to be repeated to see whether the results are similar with male patients or ones treated by therapists with a different training or experience profile too. A final and most important point is that the advantages on the one hand from the analysis being intensive (numerous observations for each subject) are limited on the other

by the small size of the sample considered (only a few subjects analysed). This prevents us from being able to generalise about the results obtained, which, however, will be useful for formulating refined hypotheses for further research on a wider scale (Stiles, 2003).

REFERENCES

Amador, X. A., & Strauss, D. H. (1990). *The Scale to Assess Unawareness of Mental Disorder (SUMD).* New York: Columbia University and New York State Psychiatric Institute.

American Psychiatric Association. (1994). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.). Washington, DC: Author.

Bach, M., de Zwaan, M., Ackard, D., Nutzinger, D. O., & Mitchell, J. E. (1994). Alexithymia: Relationship to personality disorders. *Comprehensive Psychiatry, 35*, 234–243.

Bagby, R. M., Taylor, G. J., & Parker, J.D.A. (1994). The twenty-item Toronto Alexithymia Scale. II Convergent, discriminant and concurrent validity. *Journal of Psychosomatic Research, 38*, 33–40.

Baron-Cohen, S. (1995). *Mindblindness: An essay on autism and theory of mind.* Cambridge, MA: MIT Press.

Baron-Cohen, S., Leslie, A., & Frith, U. (1985). Does the autistic child have a 'theory of mind'? *Cognition, 21*, 37–46.

Berenbaum, H. (1996). Childhood abuse, alexithymia and personality disorder. *Journal of Psychosonatic Research, 41*, 585–595.

Bruner, J. (1986). *Actual minds, possible worlds.* Cambridge, MA: Harvard University Press.

Carcione, A., Dimaggio, G., Fiore, D., Nicolò, G., Pedone, R., Procacci, M. et al. (2004, June). *Analysis of the case 306 by the Metacognition Assessment Scale (MAS).* Paper presented at the 35th Annual Meeting of the Society for Psychotherapy Research (SPR), Rome, Italy.

Carpendale, J.I.M., & Lewis, C. (2004). Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences, 27*, 79–96.

Clarkin, J. F., Yeomans, F. E., & Kernberg, O. F. (Eds.). (1999). *Psychotherapy for borderline personality.* New York: Wiley.

Dimaggio, G., & Semerari, A. (2001). Psychopathological narrative forms. *Journal of Constructivist Psychology, 14*, 1–23.

Dimaggio, G., & Semerari, A. (2004). Disorganized narratives: The psychological condition and its treatment. In L. Angus & J. McLeod (Eds.), *The handbook of narrative and psychotherapy: Practice, theory and research* (pp. 263–282). Thousand Oaks, CA: Sage.

Dimaggio, G., Semerari, A., with Carcione, A., Nicolò, G., & Procacci, M. (in press). *Psychotherapy of personality disorders: Metacognition, states of mind, and interpersonal cycles.* London: Routledge.

Endicott, J., Spitzer, R. L., Fleiss, J. L., & Cohen, J. (1976). The Global Assessment Scale (GAS): A procedure for measuring overall severity of psychiatric disturbance. *Archives of General Psychiatry, 33*, 766–771.

Falcone, M., Marraffa, M., & Carcione, A. (2003). Metarappresentazione e psicopatologia. In G. Dimaggio & A. Semerari (Eds.), *I disturbi di personalità. Modelli e trattamento: Stati mentali, metarappresentazione, cicli interpersonali* (pp. 43–76). Roma-Bari: Laterza.

Fiore, D., Pedone, R., Petrilli, D., Popolo, R., Carcione, A., & Nicolò G. (2004, June). *Interventions Metacognition Scale (I.Me.S.): A pilot study.* Paper presented at the meeting of the Society for Psychotherapy Research, Rome, Italy, June 16–19.

Fiore, D., & Semerari, A. (2003). Il disturbo borderline di personalità: il modello. In G. Dimaggio & A. Semerari (Eds.), *I disturbi di personalità. Modelli e trattamento: Stati mentali, metarappresentazione, cicli interpersonali* (pp. 107–135). Roma-Bari: Laterza.

First, M. B., Gibbon, M., Spitzer, R. L., Williams, J.B.W., & Benjamin, L. (1997). *Structured Clinical Interview for DSM-IV Axis II Disorders.* Washington, DC: American Psychiatric Press.

Flavell, J. H. (1979). Metacognition and cog-

nitive monitoring: A new area of cognitive-developmental inquiry. *American Psychologist, 34*, 906–911.

Fodor, J. A. (1983). *The modularity of mind.* Cambridge, MA: MIT Press.

Fonagy, P. (1991). Thinking about thinking: Some clinical and theoretical considerations in the treatment of a borderline patient. *International Journal of Psycho-Analysis, 72*, 639–656.

Fonagy, P., Gergely, G., Jurist, E. L., & Target, M. (2002). *Affect regulation, mentalization, and the development of the self.* New York: Other Press.

Fonagy, P., Leigh, T., Steele, M., Steele, H., Kennedy, R., Mattoon, G., et al. (1996). The relation of attachment status, psychiatric classification, and response to psychotherapy. *Journal of Consulting and Clinical Psychology, 64*, 22–31.

Fonagy, P., Moran, G. S., Steele, M., Steele, H., & Higgitt, A. C. (1991). The capacity for understanding mental states: The reflective self in parent and child and its significance for security of attachment. *Infant Mental Health Journal, 13*, 200–216.

Fonagy, P., & Target, M. (1996). Playing with reality: I. Theory of mind and the normal development of psychic reality. *International Journal of Psychoanalysis, 77*, 217–233.

Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia.* Hove, UK: Erlbaum.

German, T. P., & Leslie, A. M. (2000). Attending to and learning about mental states. In P. Mitchell & K. J. Riggs (Eds.), *Children's reasoning and the mind* (pp. 229–252). Hove, UK: Taylor & Francis.

Glick, M. J., Salvi, L. M., Stiles, W. B., & Greenberg, L. (2004, June). *When voices listen: An assimilation analysis of Lisa's problematic experiences.* Paper presented at the 35th Annual Meeting of the Society for Psychotherapy Research (SPR), Rome, Italy.

Guy, W., & Ban, T. A. (1982). *The AMDP-Sistem: Manual for the assessment and documentation of psychopathology.* Berlin: Springer-Verlag.

Hermans, H.J.M., & Kempen, H.J.K. (1993). *The dialogical self: Meaning as movement.* San Diego: Academic Press.

Hyler, S. E., & Rieder, R. O. (1987). *Personality Diagnostic Questionnaire Revised (PDQ-R). New York: Psychiatric Institute.*

Kernberg, O. F. (1975). Borderline conditions and pathological narcissism. New York: Jason Aronson.

Leslie, A. M. (1987). Pretense and representation: The origin of "theory of mind." *Psychological Review, 94*, 412–426.

Leslie, A. M. (1994). ToMM, ToBy, and agency: Core architecture and domain specificity. In L. Hirschfeld & S. Gelman (Eds.), *Mapping the mind: Domain specificity in cognition and culture.* Cambridge: Cambridge University Press.

Linehan, M. (1993). *Cognitive-behavioral treatment of borderline personality disorder.* New York: Guilford.

Liotti, G. (1999). Disorganization of attachment as a model for understanding dissociative psychopathology. In J. Solomon & C. George (Eds.), *Attachment disorganization.* New York: Guiford.

Liotti, G. (2002). The inner schema of borderline states and its correction during psychotherapy: A cognitive-evolutionary approach. *Journal of Cognitive Psychotherapy, 16*, 349–366.

Livesley, W. J. (2003). *Practical management of personality disorders.* New York: Guilford.

Lysaker, P., Carcione, A., Dimaggio, G., Johannesen, J. K., Nicolò, G., Procacci M. et al. (2005). Metacognition amidst narratives of self and illness in schizophrenia: Associations with insight, neurocognition, symptom and function. *Acta Psychiatrica Scandinavica, 112*, 64–71.

Main, M. (1991). Metacognitive knowledge, metacognitive monitoring and singular (coherent) vs multiple (incoherent) models of attachment. Findings and directions for future researches. In C. M. Parkes, J. Stevenson-Hinde, & P. Marris (Eds.), *Attachment across the life cycle.* New York: Routledge.

Modestin, J., Lötscher, K., & Erni, T. (2002). Dissociative experience and their correlates in young non-patient. *Psychology and Psychotherapy: Theory, Research and Practice, 75*, 53–64.

Neimeyer, R. A. (2000). Narrative disruptions in the construction of self. In R. A. Neimeyer & J. D. Raskin (Eds.), *Constructions of disorder: Meaning making frameworks for psychotherapy* (pp. 207–241). Washington, DC: American Psychological Association.

Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and

new findings. In G. Bower (Ed.), *The psychology of learning and motivation.* New York: Academic Press.

Nichols, S., & Stich, S. (2001). *Mindreading.* Oxford: Oxford University Press.

Nicolaou, S., Imperatore, G., Stegnano, M., Ruggeri, M., De Girolamo, G., Bellini, M., et al. (1995). *Brief Psychiatric Rating Scale (B.P.R.S.) italian version amplied (4.0). Epidemiologia e Psichiatria Sociale, 4,* 1–25.

Perris, C., & Skagerlind, L. (1998). An integrated, multilevels, metacognitive approach to the treatment of patients with a schizophrenic disorder or a severe personality disorder. In C. Perris & P. D. McGorry (Eds.), *Cognitive psychotherapy of psychotic and personality disorders.* Chichester: Wiley.

Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *The Behavioral and Brain Science, 4,* 515–526.

Ryle, A. (1997). *Cognitive analytic therapy and borderline personality disorder: The model and the method.* Chichester: Wiley.

Ryle, A., & Kerr, I. (2002). *Introducing cognitive analytic therapy: Principles and practice.* Chichester: Wiley.

Sarbin, T. R. (1986). The narrative as a root metaphor for psychology. In T. R. Sarbin (Ed.), *Narrative psychology: The storied nature of human conduct* (pp. 1–27). New York: Praeger.

Searles, H. F. (1988). *My work with borderline patients.* New York: Aronson.

Semerari, A. (2001). La disfunction metacognitiva en el Trastorno de Personalidad Borderline. *Revista de Psicoterapia, 45,* 21–37.

Semerari, A., Carcione, A., Dimaggio, G., Falcone, M., Nicolò, G., Procacci, M., et al. (2003). The evaluation of metacognitive functioning in psychotherapy: The Metacognition Assessment Scale and its applications. *Clinical Psychology and Psychotherapy, 10,* 238–261.

Semerari, A., Dimaggio, G., Nicolò, G., Procacci, M., & Carcione, A. (2005). *Understanding minds: Different functions and different disorders? The contribution of psychotherapeutic research.* Manuscript submitted for publication.

Shedler, J., & Westen, D. (2004). Dimension of personality pathology: An alternative to the Five-Factor Model. *American Journal of Psychiatry, 161,* 1743–1754.

Sperber, D. (2000). *Metarepresentations.* Oxford: Oxford Univeristy Press.

Stiles, W. B. (2003). Qualitative research: Evaluating the process and the product. In S. P. Llewelyn & P. Kennedy (Eds.), *Handbook of clinical health psychology* (pp. 477–499). London: Wiley.

Stiles, W. B., Meshot, C. M., Anderson, T. M., & Sloan, W. W., Jr. (1992). Assimilation of problematic experiences: The case of John Jones. *Psychotherapy Research, 2*(2), 81–101.

Target, M., & Fonagy, P. (1996). Playing with reality: II. The development of psychic reality from a theoretical perspective. *International Journal of Psychoanalysis, 77,* 459–479.

Taylor, G. J., Bagby, R. M., & Parker, J.D.A. (1997). *Disorders of affect regulation: Alexithymia in medical and psychiatric illness.* Cambridge: Cambridge University Press.

Tomasello, M. (1999). *The cultural origins of human cognition.* Cambridge: Harvard University Press.

Tomasello, M., & Rakoczy, H. (2003). What makes human cognition unique? From individual to shared to collective intentionality. *Mind and Language, 18,* 906–921.

Westen, D., & Shedler, J. (2000). A prototype matching approach to diagnosis personality disorders toward DSM-V. *Journal of Personality Disorders, 14,* 109–126.

Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13,* 257–270.