Affect Regulation, Mentalization, and the Development of the Self
Affect Regulation, Mentalization, and the Development of the Self

Peter Fonagy
György Gergely
Elliot L. Jurist
Mary Target
# Contents

ABOUT THE AUTHORS vii

ACKNOWLEDGMENTS xi

INTRODUCTION 1

# PART I
THEORETICAL PERSPECTIVES

CHAPTER 1
Attachment and Reflective Function: Their Role in Self-Organization 23

CHAPTER 2
Historical and Interdisciplinary Perspectives on Affects and Affect Regulation 65

CHAPTER 3
The Behavior Geneticist's Challenge to a Psychosocial Model of the Development of Mentalization 97
PART II
DEVELOPMENTAL PERSPECTIVES

CHAPTER 4
The Social Biofeedback Theory of Affect-Mirroring:
The Development of Emotional Self-Awareness
and Self-Control in Infancy 145

CHAPTER 5
The Development of an Understanding of Self and Agency 203

CHAPTER 6
"Playing with Reality": Developmental Research
and a Psychoanalytic Model for the Development of Subjectivity 253

CHAPTER 7
Marked Affect-Mirroring and the Development
of Affect-Regulative Use of Pretend Play 291

CHAPTER 8
Developmental Issues in Normal Adolescence
and Adolescent Breakdown 317

PART III
CLINICAL PERSPECTIVES

CHAPTER 9
The Roots of Borderline Personality Disorder
in Disorganized Attachment 343

CHAPTER 10
Psychic Reality in Borderline States 373

CHAPTER 11
Mentalized Affectivity in the Clinical Setting 435

EPILOGUE 469

REFERENCES 481
INDEX 549
Peter Fonagy, Ph.D., F.B.A., is Freud Memorial Professor of Psychoanalysis and Director of the Sub-Department of Clinical Health Psychology at University College London. He is Director of the Child and Family Center and of the Clinical Outcomes Research and Effectiveness Center, both at the Menninger Foundation, Kansas. He is also Director of Research at the Anna Freud Centre, London. He is a clinical psychologist and a training and supervising analyst in the British Psychoanalytical Society in child and adult analysis. He is Chair of the Research Committee and Vice-President of the International Psychoanalytic Association, and a Fellow of the British Academy. He is on the editorial board of a number of major journals, including Development and Psychopathology and the Bulletin of the Menninger Clinic. He has published over 200 chapters and articles and has authored or edited several books. His most recent books include What Works for Whom: A Critical Review of Psychotherapy Research (with A. Roth; published in 1996 by Guilford Press), Psychoanalysis on the Move: The Work of Joseph Sandler (coedited with A. Cooper and R. Wallerstein; published in 1999 by Routledge), Attachment Theory and Psychoanalysis (published in 2001 by Other Press), and Evidence-Based Child Mental Health: A Comprehensive Review of Treatment Interventions (with M. Target, D. Cottrell, J. Phillips, and Z. Kurtz; to be published by Guilford Press).
György Gergely, Ph.D., is Director of the Developmental Psychology Laboratory of the Psychology Institute of the Hungarian Academy of Sciences and Senior Lecturer at the Cognitive Developmental Doctoral Program of the Eötvös Lóránd University in Budapest. He is a clinical psychologist and is on the visiting faculty of the Max Planck Institute for Psychology in Munich, the Child and Family Center at the Menninger Clinic, the Department of Psychology at University College London, and the Department of Psychology at Berkeley. He is on the Panel of the European Cognitive Neuroscience Initiative at Trieste, Italy. He is the author of *Free Word Order and Discourse Interpretation* (published in 1991 by Academic Press of Budapest) and he serves on the editorial boards of several major journals.

Elliot L. Jurist, Ph.D., Ph.D., is a Professor in the Department of Philosophy, Hofstra University, and a Lecturer in the Department of Psychiatry, College of Physicians and Surgeons, Columbia University. He is the author of *Beyond Hegel and Nietzsche: Philosophy, Culture and Agency* (published in 2000 by MIT Press) and of numerous articles in philosophy and psychoanalysis. He is a Member of the Neuropsychiatry Service, New York Presbyterian Hospital, and a Member of the Ethics Committee of the New York State Psychiatric Institute.

Mary Target, Ph.D., is a Senior Lecturer in Psychoanalysis at University College London and an Associate Member of the British Psychoanalytical Society. She is Deputy Director of Research at the Anna Freud Centre, Member of the Curriculum and Scientific Committees, Chairman of the Research Committee of the British Psychoanalytical Society, and Chairman of the Working Party on Psychoanalytic Education of the European Psychoanalytic Federation. She is Course Organizer of the UCL Master’s Course in Psychoanalytic Theory and Academic Course Organizer of the UCL/Anna Freud Centre Doctorate in Child and Adolescent Psychotherapy. She is on the editorial board of several journals, including the *International Journal of Psychoanalysis*, and she is Joint Series Editor for psychoanalytic books at Whurr Publishers and an Associate Editor for the New Library of Psychoanalysis, Routledge. She has active research collabora-
tions in many countries in the areas of developmental psychopa-
thology and psychotherapy outcome. She is author and editor of
several books, including *Attaccamento e Funzione Riflessiva* (with
P. Fonagy, published in 2000 by Raffaello Cortina Editore of
Milan) and *The Outcomes of Psychoanalytic Treatment* (edited
with M. Leuzinger-Bohleber; to be published by Whurr).
We would like to acknowledge first of all our gratitude to our patients; some of the ideas in this book have been prompted by our clinical work, and one important yardstick in judging the ideas is whether they can help us better to understand our experiences with patients past, present, and future. We hope that our efforts toward new theoretical understanding have sometimes benefited them in turn.

This book could not and would not have been produced without the outstanding contribution of Dr. Elizabeth Allison, the Publications Editor of the Psychoanalysis Unit at UCL. Her contribution went very significantly over what might be expected from someone in this role. She effectively conceived and organized the project, as well as offering superb advice on the integration of the chapters and the dovetailing of arguments. We are sure that many gaps have remained, but the level of integration we were able to achieve is in large part due to her remarkable intellectual abilities and brilliant editorial skills. What is more, she was able to impose her editorial and intellectual discipline on all of us with charm and sensitivity. Liz, we are all immensely indebted to you.

In writing this book we have been immensely fortunate. Liz's predecessor in the Publications Editor post was Kathy Leach,
whose contribution was enormously valuable, not only in the planning phases of the project, but also in preparing for publication many of the pieces on which the present work relies. Our only regret is that despite ingenious, tactful, and sometimes desperate efforts to keep the production of this book to schedule, we were not able to reward her with a complete manuscript.

We are also grateful to our publisher, Dr. Michael Moskowitz, whose vision about a new kind of psychoanalytic press we share and are extremely pleased to be participating in. We are particularly grateful for his patience, as this book, born of work in four countries and on two continents, slowly progressed to final completion. In terms of intellectual debt, more people deserve acknowledgment than we could possibly list. The contributions of two people, however, have been clearly formative: John Watson of the University of California at Berkeley has collaborated for many years with one of us (GG), but all of us have benefited from his writings and wisdom. Even in these days of virtual workplaces, this collaboration would not have been possible without a common workplace that at least three of the four of us share. Dr. Efrain Bleiberg’s leadership of the Menninger Clinic during the time that the work reported here was undertaken has made it possible for us to work together in the Child and Family Center of the Menninger Clinic, which Dr. Bleiberg created and supervised in a truly free intellectual spirit. Others at the Child and Family Center, particularly Dr. Jon Allen and Dr. Helen Stein, have offered inspiration and advice. Two additional colleagues have been a source of inspiration and guidance: Dr. Susan Coates has followed our work over the years, as indeed we have followed hers with admiration. She has set a path in developmental clinical psychoanalysis that has been a true inspiration. Dr. Arietta Slade has inspired us not only in her research work, clinical skills, and scholarship, but the integration of all these into one personality with exceptional spirit and charm. We are grateful to her for her friendship as well as her guidance.

Earlier but somewhat different formulations of some of the ideas contained in this book have appeared, or will shortly appear, in the following publications:
Chapter 1 incorporates some material from "Attachment and reflective function: Their role in self-organization" by Peter Fonagy and Mary Target, which appeared in Development and Psychopathology, 9 (1997), 679–700.

Chapter 4 incorporates some material from "The social bio-feedback model of parental affect-mirroring" by György Gergely and John Watson, which first appeared in the International Journal of Psycho-Analysis, 77 (1996), 1181–1212.


Chapters 1 and 8 incorporate some material from "Attachment and borderline personality disorder: A theory and some evidence" by Peter Fonagy, Mary Target, and György Gergely, which first appeared in Psychiatric Clinics of North America, 23 (2000), 103–122.

The concluding chapter incorporates some material from "An interpersonal view of the infant" by Peter Fonagy and Mary Target, which first appeared in A. Hurry (Ed.), Psychoanalysis and Developmental Theory (pp. 3–31). London: Karnac, 1998.
Numerous paths come together in this book. Drawing from a wide range of sources, we ambitiously aim to address multiple audiences: research psychologists, clinical psychologists, and psychotherapists, but also developmentalists from across other disciplines. From the most general perspective, we wish to highlight the crucial importance of developmental work to psychotherapy and psychopathology. We offer an account of psychotherapy that seeks to integrate our scientific knowledge of psychological development with our experience as clinicians, working with children and adults. We believe that the interests of our patients are best served by a constant effort on the part both of individual therapists and of the profession collectively to bring about such an integration. The value of this kind of integration is by no means immediately accepted (see Green 2000; Wolff 1996), nor should it be. The psychotherapist offers clinical help, mainly though language, to people looking for help not (just) from medication, but from someone who is willing to have their minds in mind. It cannot be assumed that scientific progress in adjacent disciplines will benefit psychotherapeutic practice. For example, we can well imagine that progress in Hullian or Skinnerian learning-theory research in the 1940s and 1950s would have been of little help to psychodynamic therapists at that time. Learning theory benefited quite a different
kind of psychological therapy, less concerned with meaning and the person than with behavior and the environment.

Our work can be located within the well-established tradition of interest within psychoanalysis in developmental theory and research found among giants like Anna Freud, Melanie Klein, Mahler, Brody, Emde, Stern, and many others. A particularly inspiring example of the seamless integration of developmental and clinical thought may be found in a book by Anni Bergman (1999), Margaret Mahler's well-known collaborator. Many of the ideas in Bergman's book will be found here, perhaps cast in somewhat different perspective. All developmental orientations to psychotherapy must share many important features as they are all based on the observation of human development. At the same time, we believe that a number of the ideas discussed in this book—such as the social biofeedback theory of parental mirroring, the teleological and intentional stance, reflective function, the psychic equivalence and pretend modes of psychic reality, the alien self, mentalizing, affectivity, and, of course, affect regulation and mentalization—suggest a genuinely new direction for psychoanalysis and psychotherapy.

From another perspective, however, this book is not limited to psychoanalytic ideas and concerns. We apply a philosophy-of-mind approach in order to capture and specify the process by which infants fathom the minds of others and eventually their own minds. The notion that we fathom ourselves through others has its source in German Idealism and has been articulated further by analytic philosophers of mind (Jurist 2000). The use of philosophy of mind in this way is common in the field of social cognition. What differentiates our approach is the attention we give not just to cognition, but to affects as well. In this regard, we rely on attachment theory, which provides empirical support for the notion that an infant's sense of self emerges from the affective quality of relationship with the primary caregiver. Indeed, our work does not just borrow from attachment theory, but offers a significant reformulation of it. We shall argue that attachment is not an end in itself; rather, it exists in order to produce a representational system that has evolved, we may presume, to aid human survival. Another way to think about the contribution of this book, therefore, is as an effort to resolve
some of the historical tensions between psychoanalysis and attachment theory (Fonagy 2001).

Let us say a little more about the main theme of this work and its relation to the trio of terms found in our title. Our main focus throughout is on the development of representations of psychological states in the minds of infants, children, adolescents, and adults. Mentalization—a concept that is familiar in developmental circles—is the process by which we realize that having a mind mediates our experience of the world. Mentalization is intrinsically linked to the development of the self, to its gradually elaborated inner organization, and to its participation in human society, a network of human relationships with other beings who share this unique capacity. We have used the term "reflective function" to refer to our operationalization of the mental capacities that generate mentalization (Fonagy, Target, Steele, and Steele 1998).

Mentalization is intimately related to the development of both the agentive and the representational aspects of the self: both the "I" and the "Me" described by W. James (1890). A great deal of attention has been paid to the development of self-representation, James's "Me" or the "empirical self" (Lewis and Brooks-Gunn 1979), which encompasses the development of the set of characteristics that we believe to be true of ourselves even if this knowledge is inferred from the reactions to us from our social environment (Harter 1999). Thus, this aspect of mentalization is a concept with a rich history in both psychoanalytic theory (Fonagy 1991) and cognitive psychology (Morton and Frith 1995). However, the self as a mental agent—or, as we have referred to it elsewhere, the psychological self (Fonagy, Moran, and Target 1993; Fonagy and Target 1995)—is a relatively neglected subject of study. The relative neglect by psychologists and psychoanalysts of the developmental processes that underpin the agentive self may be seen as a residue of the traditionally powerful Cartesian doctrine of first-person authority that claims direct and infallible introspective access to intentional mind states, rather than seeing this access as a hard-won developmental acquisition. The work of Marcia Cavell (1988, 1994, 2000), among others, serves as a reminder of the limited extent to which psychoanalytic metapsychology has been able to distance
itself from Cartesian doctrine. Both psychoanalysis and developmental science have often adhered to the Cartesian tradition in their assumption that the experience of mental agency is innately given. In this book we attempt a radical break with this dominant philosophical tradition, arguing that mental agency may be more usefully seen as a developing or constructed capacity.

Developmental and philosophical studies of the representation of intentional action have revealed that the representation of intentional mind states may have a rather complex internal structure. Conscious access to these structures may be at best partial and could be totally absent. It seems to us important that we map the process by means of which the understanding of the self as a mental agent grows out of interpersonal experience, particularly primary-object relationships. Mentalization involves both a self-reflective and an interpersonal component. In combination, these provide the child with a capacity to distinguish inner from outer reality, intrapersonal mental and emotional processes from interpersonal communications. In this book we present both clinical and empirical evidence in conjunction with developmental observation to demonstrate that the baby's experience of himself as an organism with a mind or psychological self is not a genetic given.\(^1\) It is a structure that evolves from infancy through childhood, and its development critically depends upon interaction with more mature minds, who are both benign and reflective in their turn.

Our understanding of mentalization is not just as a cognitive process, but developmentally commences with the “discovery” of affects through the primary-object relationships. For this reason, we focus on the concept of “affect regulation,” which is important in many spheres of developmental theory and theories of psychopathology (e.g., Clarkin and Lenzenweger 1996). Affect regulation, the capacity to modulate affect states, is closely related to mentalization in that it plays a fundamental role in the unfolding of a sense of self and agency. In our account, affect

\(^1\)For economy and clarity, we refer to the child as “he” and to either the parent or the therapist as “she.” This is not meant to suggest any difference between male and female babies or male and female parents or therapists.
regulation is a prelude to mentalization; yet, we also believe that once mentalization has occurred, the nature of affect regulation is transformed. Here we distinguish between affect regulation as a kind of adjustment of affect states and a more sophisticated variation, where affects are used to regulate the self. The concept of “mentalized affectivity” marks a mature capacity for the regulation of affect and denotes the capacity to discover the subjective meanings of one's own affect states. Mentalized affectivity lies, we suggest, at the core of the psychotherapeutic enterprise. It represents the experiential understanding of one's feelings in a way that extends beyond intellectual understanding. It is in this realm that we encounter resistances and defenses, not just against specific emotional experiences, but against entire modes of psychological functioning: not just distortions of mental representations standing in the way of therapeutic progress, but also inhibitions of mental functioning (Fonagy, Edgcumbe, Moran, Kennedy, and Target 1993). Thus we can misunderstand what we feel, thinking that we feel one thing while truly feeling another emotion. Moreover, it is even possible that we can deprive ourselves of the entire experiential world of emotional richness. For example, the inability to envision psychological and psychosocial causation may be the consequence of the pervasive inhibition and/or developmental malformation of the psychological processes that underpin these capacities.

Our theory of affect regulation and mentalization enables us to enrich the arguments advanced by theorists such as John Bowlby about the evolutionary function of attachment. We argue that an evolutionary function of early object relations is to equip the very young child with an environment within which the understanding of mental states in others and the self can fully develop. We propose that self-reflection as well as the ability to reflect on other minds are constructed capacities that have evolved (or not) out of the earliest relationships. Since mentalization is a core aspect of human social functioning, we can infer that evolution has placed particular value on developing mental structures for interpreting interpersonal actions. Language is, of course, the major channel for symbolic interaction. Yet, in order for language to function adequately, the subjective world requires organization. Internal states must have a meaning in
order to be communicated to others and interpreted in others to
guide collaboration in work, love, and play.

It should be readily apparent that this book places consider-
able emphasis on early social experience as a primary moderator
of the mental capacities centrally involved in psychological disor-
ders and their psychotherapeutic treatment. In this respect the
work is a continuation of a psychosocial tradition that has seen
parenting practices and the early social environment as the key
to developmental psychopathology (Cicchetti and Cohen 1995; A.
Freud 1981; Masten and Braswell 1991; Rutter 1993; Sameroff
1995; Sroufe 1996). We recognize that this perspective, while
undoubtedly dominant in psychiatry and social science during
most of the past century, has come under increasing critical
scrutiny in the light of more recent evidence from behavior-
genetic adoption and twin studies (Eaves et al. 1997; Hewitt et
al. 1997). It is very probably true that past studies might have
overestimated the influence of the social environment on devel-
opment because they failed to control for genetic influences. We
note, however, that current psychiatric literature is showing
signs of a return of a naively nativist perspective, where genetic
data are viewed as indicating that a psychological level of analy-
sis cannot be productive. Consequently, early environment is
given little significance, and such significance as is accorded to it
is from a physical rather than a psychological perspective (e.g.,
Mareno and Weinberger 2000).

While it is clear that psychological principles cannot violate
the limitations set by the neurophysiological capabilities of the
systems that subserve them, this does not imply the reduction of
psychology to biology. In this book we take issue with the "genet-
icization" of human behavior. We adopt a position in line with
that of the evolutionists Dobzhansky (1972) and Gould (1987).
We consider that nature (genetics or genes) operates as a "poten-
tialist" rather than as a "determinist." Since biological determin-
ism is frequently clothed in the language of malleability and
interactionism, we note that in reality, behind these more palat-
able portrayals of biology, genes are still considered the primary
ruling force. While changeability of evolved dispositions is ac-
knownledged, these inherited traits are then seen as predomi-
nant. We suggest that this tendency can occur because of the absence of convincing data on gene–environment interaction. Yet modern genetics has provided ample room for the contention that biology has culture on a “loose” rather than a “tight leash” (Gould 1987).

We attribute the failure of behavior geneticists to find convincing evidence for gene–environment interaction to their predisposition to study the “wrong” environment: the physical rather than the psychological. We suggest that the decline of interest in subjectivity, in the mechanism that interprets the social world, is partly to blame for this gap in our knowledge. We argue that it is the manner in which the environment is experienced that acts as a filter in the expression of genotype into phenotype. The intrapsychic representational processes that underpin the agentive self are not just the consequences of both environmental and genetic effects. They may acquire additional importance as moderators of the effects of the environment upon the unfolding of genotype into phenotype. We place mentalization at the heart of this process of moderation, since it is the interpretation of the social environment rather than the physical environment that governs genetic expression.

These considerations have led us to reformulate our understanding of the impact of early social relationships on later experience. We move away from the model where an early relationship is principally seen as the generator of a template for later relationships (e.g., Bowlby 1980). Instead, we argue that early experience, no doubt via its impact on development at both psychological and neurophysiological levels, determines the “depth” to which the social environment may be processed. Suboptimal early experiences of care affect later development by undermining the individual’s capacity to process or interpret information concerning mental states that is essential for effective functioning in a stressful social world. Insecurity in attachment relationships is a signal of limitation in mentalizing skills. We find that the traditional classification of attachment patterns may be helpfully reinterpreted in this context as indication of a relatively good (secure attachment), or relatively poor (insecure attachment) capacity to manage or cope with intimate inter-
personal relationships. An absence of mentalizing capacity under stress is signaled by the disorganization of the attachment system.

So what aspect of the environment can be specified as critical to the healthy development of an agentive or psychological self? Attuned interactions with the parent (Jaffe, Beebe, Feldstein, Crown, and Jasnow 2001; Stern 1985) often involve affect-mirroring—that is, the parent’s use of facial and vocal expression to represent to the child the feelings she assumes him to have in such a way as to reassure and calm rather than intensify his emotions. We see parental affect-mirroring as instrumental in fostering the capacity for affect regulation, through the creation of a second-order representation for constitutional affect states. The image of the caregiver mirroring the internal experience of the infant comes to organize the child’s emotional experience. Thus the self is not merely open to environmental influence: it is in part constituted through its interactions with the social environment. Freud, arguably, saw infancy as a time when the self saw others as extensions of itself (e.g., Freud 1900a). Our emphasis is the reverse—we see the self as originally an extension of experience of the other.

We propose two intimately connected developmental theories. Both concern the relationship between the acquisition of an understanding of the representational nature of minds on the one hand and affect regulation on the other. The social biofeedback theory of parental affect-mirroring explores the way in which the infant’s automatic emotion expression and the caregiver’s consequent affect-reflective facial and vocal displays come to be linked in the infant’s mind through a contingency-detection mechanism identified by John Watson and colleagues (Bahrick and Watson 1985; Gergely and Watson 1996; Watson 1972, 1994). (This mechanism is described quite fully in chapter 4.) The forging of this link has two important effects: (a) The infants come to associate the control they have over their parents’ mirroring displays with the resulting improvement in their emotional state, leading, eventually, to an experience of the self as a regulating agent. (b) The establishment of a second-order representation of affect states creates the basis for affect regulation and impulse control. Affects can be manipulated and discharged
internally as well as through action; they can also be experienced as something recognizable and hence shared. Affect expressions by the parent that are not contingent on the infant's affect will undermine the appropriate labeling of internal states, which may, in turn, remain confusing, experienced as unsymbolized, and hard to regulate.

For affect-mirroring to serve as the basis of the development of a representational framework, the caregiver must somehow indicate that her display is not for real: it is not an indication of how the parent herself feels. We describe this characteristic of the parent's mirroring behavior as its "markedness." A display that is congruent with the baby's state but lacks markedness may overwhelm the infant. It is felt to be the parent's own real emotion, making the infant's experience seem contagious, or universal, and thus more dangerous. In the short term, the infant's perception of a corresponding but realistic negative emotion is likely to escalate rather than regulate the infant's state, leading to traumatization rather than containment.

Our second theory concerns the nature of subjectivity before recognition of the representational character of the human mind. We suggest that the infant's and young child's early awareness of mental states is characterized by the equation of the internal with the external. What exists in the mind must exist out there, and what exists out there must invariably also exist in the mind. Psychic equivalence, as a mode of experiencing the internal world, can cause intense distress, since the projection of fantasy to the outside world can be terrifying. The acquisition of a sense of pretend in relation to mental states is therefore essential. The repeated experience of affect-regulative mirroring helps the child to learn that feelings do not inevitably spill out into the world. The child's mental state is decoupled from physical reality. We suggest that children whose parents provide more affect-congruent contingent, and appropriately marked, mirroring displays facilitate this decoupling. In contrast, the displays of parents who, because of their own difficulties with emotion regulation, are readily overwhelmed by the infant's negative affect and produce a realistic unmarked emotion expression disrupt the development of affect regulation. A major opportunity for learning about the difference between representational and actual mental
states is lost. We argue that the equation of external and internal continues to dominate the subjective world of individuals with severe personality disorders.

Affect-mirroring can take pathological pathways, because the caregiver is overwhelmed by the negative affect generated in response to the infant's reaction and presents an overly realistic emotionally arousing display. This undermines not only the infant's possibility of creating a secondary representation, but also the sense of a boundary between self and other—an internal experience suddenly becomes external through the experience equivalent to contagion. We believe that this corresponds to the clinical characterizations of projective identification, the habitual defense particularly associated with borderline personality disorder (BPD). Sustained experience of this kind might, we argue, play an important role in establishing projective identification as the dominant form of emotional experience in the development of borderline personality. It should be noted that we use the term "borderline" in this book in the sense of a form of pathological personality organization that could be a likely factor in all serious personality disorder, rather than just the more specific DSM-IV category of personality disorder, Borderline Personality Disorder. The model advanced in this book is intended to explain borderline phenomena in many patients, not just those who meet formal diagnostic criteria. Our aim is to cast light on the nature of the mental functioning of a larger group of patients—perhaps those who fit Otto Kernberg's (1967) classical description of borderline personality organization—whose thinking and emotional experience often becomes massively disorganized and regressed in the clinical setting, evoking intense feelings in the therapist (e.g., Rey 1979). The emotional lability that is a hallmark of these patients is generally regarded to be a clue to the nature of their disturbance. Their therapy rarely takes place without dramatic enactments—their own and sometimes their therapists'—and their intense dependence on the therapist increases the difficulties of the treatment process.

A second type of deviant mirroring structure is one that we feel predisposes to narcissistic personality disorder rather than to borderline states. When affect-mirroring is appropriately
marked but is noncontingent, in that the infant's emotion is misperceived by the caregiver, the baby will still feel the mirrored affect display to map onto his primary emotion state. However, as this mirrored state is incongruent with the infant's actual feelings, the secondary representation created will be distorted. The infant will mislabel the primary, constitutional emotional state. The self-representation will not have strong ties to the underlying emotional state. The individual may convey an impression of reality, but as the constitutional state has not been recognized by the caregiver, the self will feel empty because it reflects the activation of secondary representations of affect that lack the corresponding connections within the constitutional self.\(^2\) Only when psychotherapy generates mentalized affectivity will this fault line in the psychological self be bridged.

We have attempted to specify in greater detail the psychological mechanism that underpins the processes by which parental affect-mirroring deviates from its normal course in terms of our concept of the *alien self*. In general we might say that the self as agent arises out of the infant's perception of his presumed intentionality in the mind of the caregiver. Where parental caregiving is extremely insensitive and misattuned, we assume that a fault is created in the construction of the psychological self. We follow Winnicott's (1967) suggestion that the infant, failing to find *himself* in the mother's mind, finds the mother instead. The infant is forced to internalize the representation of the object's state of mind as a core part of himself. But in such cases the internalized other remains *alien* and unconnected to the structures of the constitutional self. In the case of chronically insensitive or misattuned caregiving, a fault is created in the construction of the self, whereby the infant is forced to internalize the representation of the object's state of mind as a core part of himself.

In early development this "alien self" is dealt with by externalization; as mentalization develops, it can be increasingly

\(^2\) By "constitutional self" we mean an individual's biologically determined experiences associated with emotional expressivity as well as temperamental manifestations of affect.
woven into the self, creating an illusion of cohesion. Thus, the young child showing disorganized attachment will frequently control and manipulate the parent's behavior. This is part of a projective identificatory process whereby he achieves his need to experience his self as coherent, and the alien part of his self-structure as outside his mind, by perceiving these elements within other selves, normally within a parent. The disorganization of the self disorganizes attachment relationships by creating a constant need for this projective identification—the externalization of the alien self—in any attachment relationship.

The alien self is present in all of us, because transient neglect is part of ordinary caregiving; with the development of mentalization, and given a mid-childhood environment relatively free of trauma, the gaps in the self corresponding to noncontingent parenting are covered over by the self-narratives that the reasonably functioning mind can create. The alien self is mostly pernicious when later experiences of trauma in the family or the peer group force the child to dissociate from pain by using the alien self to identify with the aggressor. In these cases the gaps come to be colonized by the image of the aggressor, and the child comes to experience himself as destructive and, in the extreme, monstrous. Thus we see inadequate early parenting as creating a vulnerability that may become destructive of development and seriously pathogenic if later experience is unfavorable, either in terms of failing to facilitate the later development of mentalization or explicitly calling for the use of the faults in self-development for defensive purposes. These factors interact, and the likelihood of surviving psychological assault improves if mentalizing is freely available to interpret the perpetrator's behavior (Fonagy, Steele, Steele, Higgitt, and Target 1994). However, brutalization in the context of attachment relationships generates intense shame. This, if coupled with a history of neglect and a consequent weakness in mentalization, becomes a likely trigger for violence against the self or others, because of the intensity of the humiliation experienced when the trauma cannot be processed and attenuated via mentalization. Unmentalized shame, which remains unmediated by any sense of distance between feelings and objective realities, is then experienced as
the destruction of the self. We have called it "ego-destructive shame." The use of the alien, dissociated part of the self to contain the image of the aggressor and the unthinkable affect generated by their abusive actions is a survival strategy for many individuals, some of whom we try to describe in this book.

The defensive use of the alien part of the self is deeply pathogenic, although initially adaptive. In our view it marks the development of serious personality problems. It involves three important changes: (a) a further repudiation of mentalization, at least in attachment contexts, (b) disruption of the psychological self by the emergence of a torturing other within the self, and (c) vital dependence on the physical presence of the other as a vehicle for externalization. These features, in combination, account for many aspects of disordered functioning in borderline patients. Abused and traumatized individuals, unable to conceive of the mental states that could explain the actions of the abuser, voluntarily and defensively sacrifice their thinking about internal states. While able to think of mental states in self and other in the context of ordinary social relationships, they inevitably become conflicted and entangled once a relationship becomes emotionally intense, organized by mental structures that are involved in attachment relationships. The abandonment of mentalizing leaves them with an internal reality that is dominated by psychic equivalence. These individuals, like all other patients, organize the therapeutic relationship to conform to their unconscious expectations, except that for them these expectations have the full force of reality and there is no sense of alternative perspectives. The inability to think about mental states removes the possibility of "narrative smoothing" of the basic gaps in the self-structure, and the alien self emerges in a manner much clearer for the therapist to see and experience. Splitting becomes a dominant defense and projective identification—the externalization of the alien self—an essential part of survival. The vehicles for the projective identification must be present for the process of externalization to function, and total dependency on these individuals turns into a dominant theme.

Our reformulation of the significance of the early social environment has important clinical implications. Psychotherapy
with individuals whose early experiences have led to a compromised mentalizing capacity should be focused on helping them to build this interpersonal interpretive capacity. One way of conceptualizing the entire psychotherapeutic enterprise may be as an activity that is specifically focused on the rehabilitation of this function. The work of many previous theoreticians can be reread in the light of our reformulation. Throughout the book we often point out that the ideas being advanced here were anticipated in the work of great psychoanalytic theoreticians, not least those in the object-relations tradition. For example, we make many connections with Bion’s (1959) thinking about containment. The interpersonal interpretive stance of the human brain is an overarching biological construct (Bogdan 2001). Its limits are as yet unexplored. Our concerns in this book are principally with reflective function and mentalization—capacities that are likely to turn out to be specific instances of this more general capacity.

An important aim of psychotherapy, then, if not its central aim, is the extension of mentalization. With some patients, particularly those at the borderline end of the spectrum, the therapist’s task may be considered to be similar to that of the parent who intuitively engages with the child’s world of psychic equivalence to emphasize its representational character. Integrating the concrete and dissociated (pretend) modes of functioning can only be achieved through focused work on the patient’s current experience of the transference. Some enactments on the part of both therapist and patient are an inevitable part of this work, since the patient cannot remain psychologically close to the therapist without externalizing the alien parts of the self. It is at these moments, when the therapist is enacting the split-off part of the patient’s experience, that the patient’s true self may be most accurately observed. Sadly, almost invariably the therapist will have extreme difficulty in communicating insight and understanding at such times. It is far more likely, in the middle of such turmoil, that the therapist’s rage or terror or both will obscure her vision of the patient. Nevertheless, persistence and a determined focus on understanding the moment-to-moment changes in the patient’s experience usually pays off, and the psychotherapeutic treatment of even quite severely impaired patients can be
surprisingly effective (Bateman and Fonagy 1999) and lasting (Bateman and Fonagy 2001).

Not all borderline patients are as severely disturbed as the above description implies, although most have features contained in it. If the caregiver has presented affect displays to the infant that correspond poorly to the child's constitutional self-state, affect regulation will be based on second-order representations of affects that do not have strong connections with constitutional self-states. Inaccurate mirroring by the caregiver will lead to emotional experiences that cannot be experienced as "true." Consequently, the secondary representational structures will not provide a means for accessing and attributing emotion states to the self. So words will not form the basis for the child's emerging ability to think about his or her real thoughts and wishes. The aim of psychotherapy for these individuals is to regenerate the connection between the consciousness of an affect state and its experience at the constitutional level. We have labeled this "mentalized affectivity"—a term intended to indicate the capacity to connect to the meaning of one's emotions. The clinical emphasis on experiential understanding of one's feelings in a way that ensures "meaningfulness" is crucial because it serves to establish congruent connections between primary and secondary affect-representational structures. The focus on emotion ensures that the secondary representational structures used to think about or to reflect on affect are reconnected and that misconnections where displayed affect was tied to a different nonconscious affect state are corrected.

This book is divided into three parts. The first part (chapters 1–3) is theoretical; the second part (chapters 4–8) is developmental; the third part (chapters 9–11) is clinical. The book concludes with an Epilogue. The tripartite division is primarily a matter of convenience; it is not meant to convey that the components of the content are separate. The reader will find, for example, that in the theoretical part there is considerable discussion of developmental material; that in the developmental part there is important theoretical material on the self as well as the introduction of some clinical material. This is deliberate. Throughout this work, we attempt to integrate theory, development, and the clinical sphere. We are committed to forging a theory that is inspired by
clinical observation and rooted in the findings of systematic research. Our aim is to bring both theory and research to bear on clinical practice.

The content of the chapters is as follows: Chapter 1 offers an overview of the major themes to be covered in the book. We survey the links between attachment theory, studies of early social development, and studies of cognitive development that focus on one approach to mentalization—the study of the acquisition of a "theory of mind." We sketch an outline (to be filled in later) of what we know about the developmental course of self-organization and the developmental deviations that can result from an adverse early environment, later trauma, or both. We introduce the notion of mentalization as a potential mediator of psychosocial risk.

In chapter 2, we consider broad intellectual trends in theories of affect, across a range of disciplines. We suggest that the dialectic between those who see affect as inevitably disorganizing and external to cognitive influence and those who attempt to integrate affect as a form of cognitive activity may be addressed through the concept of mentalization and—more narrowly—mentalized affectivity. We introduce a distinction between first- and second-order representation of affect and argue that this distinction is consistent with data from psychology, neurophysiology, and psychoanalytic theories.

In chapter 3, we address the challenge posed by recent arguments about the relative importance of genes and the early environment to a mainly psychosocial model of personality development. It may seem foolhardy to advance a model in which the capacity for mentalization is rooted in an attachment relationship, at a time when social cognitive capacities are increasingly seen as being genetically determined rather than as products of the early environment. We counter this challenge by arguing that the human capacity for interpreting the social environment is a key moderating influence in the expression of genes in phenotypes. Mentalization and the family of capacities to which it belongs (deriving from the "interpersonal interpretive mechanism") can be either facilitated or retarded by attachment experiences and can play a key role in creating an experienced environment. As part of this argument, we propose a reorienta-
tion of attachment theory, from an emphasis on templates for relationships, fixed in early infancy, to a model that views attachment as the context provided by evolution for the development of interpersonal understanding.

Chapter 4 presents in detail our developmental model of how an interpersonal interpretive mechanism for affect might evolve intersubjectively, in the context of the infant-caregiver relationship. This is a parsimonious model. Traditionally, psychoanalytic developmental models have made extravagant assumptions about infants' capacities. This was one of the major foci of the Controversial Discussions between Anna Freud and Melanie Klein (King and Steiner 1991). In contrast, our more parsimonious model is based on Gergely and Watson's (1996) Social Biofeedback Theory, which is, in turn, rooted in Watson's work on the developmental functions of the human infant's sensitivity to contingent relationships between his proprioception (of intentional movement) and the external world. We argue that the internal representation of affect develops around the mother's mirroring of the infant's expressions of affect. The caregiver's mirroring display is internalized and comes to represent an internal state, but it can do so only in certain conditions, which include sufficient attunement, together with signaling to the infant that the affect the caregiver is expressing is not her own but the child's. This account will help us to understand a range of severe pathologies.

In chapter 5, we review developmental research concerning the development of the self as agent rather than as representation. We consider a five-phase model of gradually increasing sophistication concerning the mental world, both of the self and of others. In this chapter we confront the critical issue of intersubjectivity, a fraught notion within developmental theory. We argue that interpersonal awareness is a relatively late developmental acquisition. Assuming a developmental progression toward full interpersonal awareness offers us a rich range of developmental explanations for psychological disturbances. In general, we maintain throughout this volume that personality disturbances often reflect the reemergence of rudimentary forms of interpersonal awareness—an argument that is incompatible with the assumption of an inherent (primary) intersubjectivity.
In chapter 6, we attempt to apply the findings from early developmental research more directly to construct a psychoanalytic model of the growth of subjectivity. Using clinical as well as empirical evidence, we identify two incompatible and probably alternating ways in which young children treat their inner experiences. The mental world can feel either totally real or totally unreal to a young child, and we suggest that playing with reality, making the real unreal and vice versa, is the principal avenue for the development of mentalization. We introduce the cases of two young children, both treated in psychoanalytic therapy. "Rebecca" was the loved child of young single mother, a child who learned in therapy to play with her painful image of her father. This idea had initially been too real to be played with, and the feelings it generated obstructed her otherwise normal development.

In chapter 7, we introduce "Mat," a far more deprived child. Mat's experience of inadequate mirroring left him with uncontainable and uncontainable affect, which he experienced as part of his bodily self; in an effort to create an illusion of self-control. His encopresis and his quasi-autistic features both responded to a playful therapeutic stance.

In chapter 8, we consider the special case of adolescence, a time of high risk for the emergence of serious disturbance. We argue that this is the consequence of an increased load on the mental world of the adolescent, brought on by the sudden leap in cognitive complexity, simultaneously with pressure for separation from the primary attachment figures. We consider two cases, "Tony" and "Glen." Both were challenged by the sharply increased complexity of their interpersonal worlds, but once again they followed quite different paths. We maintain that this reflected critical differences in their early experience, which equipped one with a far more robust mentalizing capacity than the other.

In the two following chapters, we discuss severe personality disorder in adults, which we see as the developmental continuation of poorly established mentalized subjectivity.

In chapter 9, we outline the theory that borderline personality disorder results from the inhibition of mentalizing due to maltreatment in childhood. We expect that this kind of inhibition is
more likely to occur if, because of inadequate early parenting, the precursors of mentalization were not firmly established. Interpersonal interpretive processes that developmentally antedate mentalization appear to govern the behavior of individuals with BPD, at least within attachment relationships. In chapter 10, we link the failure of mentalization to a distortion in the structuring of the self; failures of early attunement are expected to cause disorganization within the structure of the self-representation. Combined with trauma, these limitations in the integrity of the self can cause massive disruption in relationships. Two cases are presented. In the first case, "Emma," lack of mentalization was evident in her treatment of her own body—her suicidality and mismanagement of her chronic physical illness. The disorganization of her self-structure, when externalized onto her body, led to a profound distortion of the link to her physical experiences. The second case, "Henrietta," had a history of violence against a lover, as well as self-destructiveness. Her functioning at a prementalistic level profoundly colored her analytic relationship. In this chapter we also consider the implications of a disorganized self-structure, with very limited capacity for mentalization, for our understanding of certain transference and countertransference phenomena.

Finally, in chapter 11, we consider four more therapeutic encounters. Each illustrates in a different way the centrality of mentalized affectivity in the course of psychoanalytic therapy. The aim of this chapter is to illustrate the way misconnections between primary and second-order representation of affect can generate disturbances of self-development, which were earlier argued to be the consequences of inaccurate mirroring. To address such pathologies, the therapy must simultaneously activate primary and secondary representations of affect. The cases illustrate the importance of the integration of affect and cognition in the practice of psychoanalytic psychotherapy. Through developing second-order representations for emotional states we can come to experience our subjectivity in a way that bridges two fundamental needs in human beings. On the one hand, we need to maintain a sense of the internal equilibrium between mental forces—a struggle that was at the heart of Freud's journey of discovery. On the other, we need to be integrated into the social
world, in which we must respect the separateness of other minds, while being able also to build flexible bridges across that separateness, to create close emotional and working relationships. Therapy only works when both intrapsychic and interpersonal aims are achieved, and this is where we believe that mentalizing affectivity plays a vital role.

The Epilogue explores some of the implications of our ideas for psychopathology and psychotherapy in general.


Alessandri, S. M. (1992). Mother–child interactional correlates of mal-
References

Auerbach, J. S. (1993). The origins of narcissism and narcissistic


References


Nezworski (Eds.), *Clinical Implications of Attachment* (pp. 41–94). Hillsdale, NJ: Lawrence Erlbaum.


References


Cassidy, J., Marvin, R. S., and The MacArthur Working Group on


References


the serotonin transporter gene. *Psychological Medicine, 29* (3), 735–739.


Drucker, J. (1975). Toddler play: Some comments on its functions in
the developmental process. *Psychoanalysis and Contemporary Science, 4*, 479–527.


Field, T., Healy, B., Goldstein, S., Perry, S., Bendell, D., Schanberg, S.,
References


Fonagy, P., Steele, M., Steele, H., Leigh, T., Kennedy, R., Mattoon, G.,


Gergely, G. (1997). Margaret Mahler's developmental theory reconsider-


References


Gopnik, A., and Graf, P. (1988). Knowing how you know: Young chil-
**References**

dren's ability to identify and remember the sources of their beliefs. *Child Development, 59*, 1366–1371.


Legerstee, M., and Varghese, J. (2001). The role of maternal affect
References

References


and P. Shaver (Eds.), *Handbook of Attachment*. New York: Guilford Press.


Main, M., and Hesse, E. (1992). Disorganized/disoriented infant behavior in the Strange Situation, lapses in the monitoring of reasoning and discourse during the parent's Adult Attachment Interview, and dissociative states. In: M. Ammaniti and D. Stern (Eds.), *Atta-
References


Marenco, S., and Weinberger, D. R. (2000). The neurodevelopmental...


References


van der Kolk, B. (1994). The body keeps the score: Memory and the


Wolff, P. H. (1996). The irrelevance of infant observations for psycho-


