Bad Blood Revisited: Attachment and Psychoanalysis, 2014

Peter Fonagy
Research Department of Clinical, Educational and Health Psychology, University College London, UK

Chloe Campbell
Research Department of Clinical, Educational and Health Psychology, University College London, UK
Introduction

There is bad blood between psychoanalysis and attachment theory. As with many family feuds, it is hard to identify where the problem began. (Fonagy, 2001, p.1)

Thus began an attempt, *Attachment Theory and Psychoanalysis*, written 15 years ago, to trace the relationship between attachment and psychoanalysis, disentangle conceptual similarities and differences, and point to areas of convergence that were then emerging in the field. Here, we will attempt a conceptual progress report on how thinking has moved on since that time, with a particular emphasis on the latest developments in the theory of mentalizing as one of the current strands of thought with a bearing on the intellectual rapprochement between attachment and psychoanalysis; we find that there are new kinds of difficulties facing the field which require a reconsideration of where future directions may lie – for both attachment and psychoanalysis.

To place attachment within the context of Thomas Kuhn’s philosophy of science, described in *The Structure of Scientific Revolutions* (Kuhn, 1996), we suggest that attachment might be on the verge of a paradigm shift into a phase of ‘science revolution’. It experienced its ‘pre-science’ in the 1950s and 1960s, when attachment was first formulated by Bowlby and colleagues such as Mary Ainsworth; in the last forty years or so, attachment work has enjoyed a period of ‘normal science’, in which a slew of major research breakthroughs have taken place in a context of well-established methodological and theoretical parameters. But, as increasingly contradictory and less-clear cut data surrounding attachment emerges, which demands less reflexive theoretical interpretation and which starts to question how sufficient or accurate existing predominant research measures can be, we are possibly now approaching another paradigm shift, into the era of ‘revolutionary science.’ In this
paper, we would like to tentatively anticipate this shift in attachment thinking, and point to the bearing this might have on the psychoanalytic project and its relationship with attachment.

**Lines of convergence between attachment and psychoanalysis**

There have been several trends within modern psychoanalysis that have paved the way for a greater engagement with attachment. This is an area that has been covered in depth elsewhere, and does not need detailed reiteration here (Eagle, 2013) (Fonagy et al., 2008, Fonagy, 2001). In brief terms, we can see that psychoanalysis has become more pluralistic and accepting of differences (Holmes, 2009, Fonagy and Target, 2007a). Bowlby formulated attachment theory in riposte to the Kleinian and classical Freudian approaches as they existed in the mid-twentieth century. In parallel with the emergence of attachment theory, psychoanalysts were also independently suggesting corrections or counterarguments to the Kleinian emphasis on endogenous infant fantasy, or the Freudian emphasis on the pre-eminence of innate drives. Just as attachment theory has undergone refinement and evolution, so did psychoanalysis, and many of these changes should allow for greater connections between the two schools of thought. It began with the work of Donald Winnicott and Ronald Fairbairn in the 1950s and 1960s and continued with Heinz Kohut’s selfpsychology in the 1970s, and to some degree even took over the bastions of object relations sensitive ego-psychology (e.g. Kernberg, 1991). The ‘relational movement’ emanating from the White Institute continued to have a mixed attitude to attachment theory with some degree of acceptance of the major findings (Mitchell and Aron, 1999, Mitchell, 2000) but a profound rejection of its non-hermeneutic, positivist epistemology (Hoffman, 2009).

In all, however, it would be fair to say that the growing dominance of a relational and relationship-focused emphasis in ‘modern’ psychoanalysis over recent decades has resulted in an increasing implicit interest in objective observations concerning the formative nature of the child’s social environment, and the emergence of object relations theory as the ‘lingua franca’ of modern psychoanalysis has played a pivotal role in this context (Epstein, 2010, Brown, 2010, Aron and
Leichich, 2011). Concern with the social environment was driven by an increasing interest in infant development as a legitimate way of explaining differences in adult behaviour. For example, if we accept Fairbairn’s assertion that people are fundamentally driven by relationships and their need for them, it becomes self-evident that the pursuit of relationships is not a secondary by-product of the primary drives for gratification described by Freud but rather a drive to achieve some kind of Weberian ideal type\(^1\). Consequently, an infant’s psyche is shaped by a hitherto only vaguely understood process of aggregation of early relationship experiences – a domain in which different branches of psychoanalytic thinking tend to highlight specific categories of dyadic (infant-caregiver) subjective experience as dominant in ensuring ‘good enough’ psychic maturation, e.g. mirroring in Kohut (1972), holding in Winnicott (1972) and containment in Bion (1962), etc. There are clear congruencies here, then, with attachment theory, which could be seen as simply pointing to another kind of developmentally significant dyadic configuration linked ultimately to sensitive responding (Ainsworth et al., 1978). As the object relations model moved to replace ego psychology as a dominant international psychoanalytic paradigm, so attachment theory’s emphasis on the innate need for a relationship as both the driver of intentional (thought and feeling driven) action and the primary underpinning for mental development came, at least implicitly (if not consciously), to be accepted by a majority.

Similarly, attachment theory has undergone changes that have potentially opened common ground with the psychoanalytic approach. In particular, Bowlby’s formulation of the mechanisms by which ‘maternal deprivation’ might translate into manifold forms of psychic distress only became more fully and subtly elaborated in the second volume of the trilogy, *Attachment, Separation* and *Loss* via the concept of the representational, or working, model (Bowlby, 1969, Bowlby, 1973, 1980).

---

\(^1\) Weber’s (WEBER, M. 1923. *Economy and society*, New York, NY, Bedminster Press, 1968. concept is helpful not just at the sociological level but also in the consideration of the organisation of human cognition. “An ideal type is formed by the one-sided accentuation of one or more points of view and by the synthesis of a great many diffuse, discrete, more or less present and occasionally absent concrete individual phenomena, which are arranged according to those one-sidedly emphasized viewpoints into a unified analytical construct.”
Bowlby, 1980). While Bowlby maintained his evolutionary and ethological perspective, this later thinking led to a movement away from conceiving attachment as simply analogous with imprinting (which encapsulated the mindlessly mechanistic quality that many psychoanalysts attributed to attachment theory), and towards “the level of representation” (Main et al., 1985). As exemplified by an entire industry of inquiry based on Mary Main’s Adult Attachment Interview, attachment theorists felt enabled to examine attachment in older children and adults. This led to the reconceptualization of individual differences in attachment organization as individual differences in the structure and functioning of relational mental representations (in effect ‘the self in relation’).

What a psychoanalyst such as Kohut might link with integrity of the self, Main would describe as the primary determinants of the state of mind in relation to attachment. The study of attachment related narratives was accepted as providing a window on the internal working model (Bretherton and Munholland, 2008). This once again brought attachment theorists into a common domain of theoretical and clinical discourse with psychoanalysts writing from an object relations perspective. In essence, the ‘operationalization’ of IWMs, as ‘structured processes’ (Main et al., 1985, Main et al., 2011), allowed attachment theory to accommodate greater human complexity, fluidity and subjectivity than was expressed in the reflexive, Pavlovian behaviours that early psychoanalytic critiques attributed to attachment-based interpretations.

**Ongoing criticisms of attachment theory within psychoanalysis**

This account of a happy intellectual rapprochement between attachment and contemporary psychoanalysis is only a partial telling of recent history (Appelbaum, 2011, Eagle, 2013). Disquiet about the implications of attachment research has remained. Much of the recent criticism of attachment thinking and its work has arisen from concern about the empirical research method, as some of psychoanalysis’s most talented contemporary voices have expressed. Adam Phillips wrote: "I am personally not at all interested in research, I think there is something compliant and servile about believing you have to meet the dominant criteria, and I don't think psychoanalysts should
have bought into the scientific model with such eagerness. I don't think psychoanalysis is a science or should aspire to be one." (Rustin, 2012)

Other eminent psychoanalysts have remained sceptical of the role that infant research can usefully have upon the psychoanalytic project. In 2005, the late Andre Green eloquently bemoaned the drawing together of conflicting theories as generating a mere ‘illusion of common ground’ (Green, 2005). Similarly, Irwin Hoffman has criticized attempts to integrate psychoanalytic work with the pursuit of research based on controlled studies, at the expense of clinical case studies. Arguing that this work is damaging to the analytic process – both in terms of clinical practice and our understanding of psychoanalysis itself – he defends the uniquely psychoanalytic voice which ‘stands up for the full richness, complexity, and mystery of each moment of human experience and for its manifold unrealized potentials.’ (Hoffman, 2009, p.1065)

As expressed previously by Fonagy and Target (Fonagy and Target, 2007c, p.415): attachment theory may indeed be limited from a psychoanalytic perspective in that ‘it sidesteps sexuality [Zamanian, 2011, Fonagy and Target, 2002]; sees aggression as secondary to more fundamental motivations [Fonagy et al., 1993]; arguably offers mechanistic models of conflict [Fonagy and Target, 2000]; is moot on unconscious fantasy [Fonagy and Allison, 2011]; is reductionist in its focus on a handful of empirical paradigms (e.g., the Strange Situation and the Adult Attachment Interview) that provide broad classifications which lose the subtlety and detail of the original material [Fonagy and Target, 2007b]; and offers a limited framework for clinical work’ (Slade, 2000).

These issues, however, deserve more detailed consideration: closer scrutiny does identify profound differences of approach, but it also reveals misunderstandings about attachment that have arisen from superficial readings. For example, given Bowlby’s interest in unconscious defences against memories of traumatic separation and loss, and the detailed work of other attachment theorists on many other defences which unconsciously structure the developing personality and
capacities for relating, it would be wrong to claim that attachment theory does not concern itself with what psychoanalysis designates as its defining domain of interest, ‘the dynamic unconscious’. However, what attachment theory has so far undoubtedly lacked has been a nuanced interest in the multi-layered nature of human subjective experience, moving from consciously experienced expectations to the more subtle influences of internal working models on cognition, and contexts where the imperative to maintain emotion regulation overrides the demand for the veridicity of mental representations. To state it bluntly and perhaps somewhat unfairly, there is little attention paid by attachment theorists to the qualitative differences between conscious, preconscious and unconscious experience, which is essential to the understanding of the mind as at war with itself. Behind the multi-layered mind of classical psychoanalytic study is an implicit psychobiological approach, where psychic contents – which are assumed to be most formatively defended against – are rooted in the conflict between the inherently (or secondarily) self-destructive nature of human minds and the need for social collaboration and relating. This, in essence, is an evolutionary model where phylogensis and ontogenesis posit potentially incompatible layers of strategy laid down in developmental and evolutionary time (Werner and Kaplan, 1963, Werner, 1948). A sole focus on the development of the self and the self in relation to another, which is part of the attachment theory and some of modern psychoanalysis, struggles to justify the multi-layered character of subjectivity exactly because it eschews this arguably outdated model rooted in nineteenth-century embryology.

It is not that attachment theory lacks speculative evolutionary psychology. The major difference between Bowlby’s thinking and Freud’s was in Bowlby’s perception of the human emotional need for others as innate, universal, and evolutionarily driven. Freud, on the other hand, saw the specificities and dark complexities behind the impulses involved in human relationships, allowing for an exploration of the mind – with its fantasies, dreams, its irrationality and its trickeries –which more readily allows for the dislikeable and contradictory in human emotions. Without such complexity, whether justified or not by empirical observation, some of the nuances of dynamic
psychology are inevitably and incontrovertibly lost. Without capturing more of human complexity, the appeal of attachment theory remained limited.

Against (arguably unnecessary) complexity, attachment theory offered parsimony based on empiricism. Yet the paradigm-bounded character of attachment research served to compound its reductionist image. The first decades of studies on attachment were strongly influenced by the Strange Situation research protocol and its corollary in adulthood, the Adult Attachment Interview. There has been a particular emphasis on assessing the stability and enduring nature of attachment status from infancy to adulthood (Aikins et al., 2009, Pinquart et al., 2013, Raby et al., 2013, Moss et al., 2005), and on the connection between parental attachment and the transmission of security or insecurity across generations\(^2\) through parenting styles (Botbol, 2010, Atkinson et al., 2005, Belsky, 2006a, Bernier et al., 2014). The success of the Strange Situation as a repeatable, easily conducted protocol, has led to its predominance as a measure within the attachment field but detracted from the clinical subtlety of attachment work. Interestingly, Mary Ainsworth, who devised the Strange Situation, herself expressed some dismay that its popularity distracted from the other research methods she used in investigating maternal sensitivity (Bretherton, 2003). Some of her own work that Ainsworth considered the most significant was her more time-consuming observational studies of parenting; her study in Baltimore in the early 1960s, for example, involved monthly home visits, each of 4 hours’ duration. During this time the mother was encouraged to interact naturally with her infant and go about her day normally. Mother-infant interactions during feeding, close physical contact, crying and playing were carefully noted. Particularly influential was the finding that a mother’s appropriate responsiveness to crying in the early months predicted less crying later on in

---

the infant’s first year. Ainsworth’s carefully modulated and painstakingly executed home observation practice was less easy to replicate than the Strange Situation, and Ainsworth herself commented: ‘I have been quite disappointed that so many attachment researchers have gone on to do research with the Strange Situation rather than look at what happens in the home or other natural settings – like I said before – it marks a turning away from field observations and I don’t think it’s wise’ (Bretherton, 2003, p.325).

Similarly, the emphasis on the Strange Situation protocol and the Adult Attachment Interview may also detract from some of the complexities of Bowlby’s thinking. Mary Main’s genius in seeing the communicational quality of attachment narratives as the legacy of early attachment and as indications of current states of mind in relation to attachment concerns has yielded discovery upon discovery and its extraordinary value cannot be denied (Main et al., 2011). Yet the reduction of the internal working model to Grice’s maxims of conversation (Grice, 1975) takes the focus away from narrative content and moves scientific discourse into the domain of mechanism, and has perhaps resulted in a reductionist over-simplification of the concept of the IWM. Bretherton’s work on IWMs has shown how internal, symbolic processes are not necessarily ignored in attachment theory; the IWM was described as a representation of the self in metaphorical conversation with the other (Bretherton and Munholland, 1999). The internal, psychic power of the IWM as expressed by Bowlby was partly disguised by his antipathy towards the psychoanalytic tendency to focus on internal fantasy at the expense of real-life experience. Bowlby’s emphasis on how early environmental experiences mould the IWM does not in fact detract from the richness and imaginative complexity of the IWM that each child devises from their experiences.

**Remaining challenges**

The fields of both attachment and psychoanalysis are now confronting the challenging and complex reality presented by empirical data, and the daunting fact of data’s inability to support simple schema. Developments within attachment research have suggested a need for greater nuance and
refinement in relation to attachment thinking, which we will consider here as having potential bearing on future directions in attachment work and its relationship with the wider psychoanalytic project. It is interesting to note that the data’s recalcitrance when it comes to respecting neat predictions mirrors some of the very concerns that psychoanalysts have raised about the erosion of human complexity in the pursuit of schema and mechanisms by developmental researchers.

**Beyond the normal science of attachment**

Recent research has increasingly pointed to the limitations of the evidence linking the early childrearing environment to later outcomes and the limited power of prediction which the observation of early relationships offered, indicating that there is much further work to be done on the mechanisms of attachment and the processes involved in tracing its impact in later life (Fearon et al., 2014, Luyten, in press). The independence of attachment from genetic influences may well turn out to be characteristic only of infantile and early childhood attachment types, and to be subject to the unfolding of genetic influences in the adolescent years (Fearon et al., 2014). Similarly, other recent work has indicated the fluctuating role of infant attachment and genes across the life trajectory (Raby et al., 2013), suggesting that an individual’s attachment style changes in adaptation to genes and the social environment (Roisman et al., 2007, Pinquart et al., 2013). Meanwhile, a major meta-analysis by De Wolff and Van IJzendoorn has found that maternal sensitivity (defined as responsiveness to the baby’s signals) – previously considered the royal road for the transmission of attachment security in infancy – is only one factor, albeit an important one (De Wolff and van IJzendoorn, 1997). Other factors, such as stimulation and positive attitude also showed a relationship with attachment security, as did wider contextual factors, such as socioeconomic environment. Similarly, there are indications that infants may inherit differential genetic susceptibility to the emotional environment (Belsky, 2006b) making genetics (perhaps even single genes) critical moderators of environmental influence. The accumulation of evidence uncovering

---

mediating and moderating factors in the relationship between early life and developmental outcomes in one sense suggests a simple conclusion: the psychoanalytic criticism was right all along, *attachment is not everything.*

As attachment theorists we may be reluctant to jettison the explanatory framework of attachment for understanding early emotional development. But we recognise that there is a need to account for the increasingly complex and fluctuating picture of the influence of attachment processes and their predictive power over the life cycle. While research has maintained that attachment is an evolutionary-driven instinct for both the physical preservation and the emotional and cognitive development of infants, research has also indicated that the ways in which attachment ties are forged between infants and primary caregivers are highly contingent on environmental factors. This should not surprise us if we maintain an evolutionary perspective on attachment – a process as cognitively and emotionally charged as attachment would cease to adaptive if it were not responsive to wider cognitive and emotional stimuli, and indeed as Belsky and others have postulated, this in itself may account for the preponderance of different attachment styles across any population (Belsky, 2006b, Mikulincer and Shaver, 2007).

Recent interesting attachment research has reclaimed the detailed, intricate observational work originally undertaken by Ainsworth – work that has more parallels with the infant observation psychodynamic tradition, for example Beebe and Lachmann’s close, microanalytic work on early parent-infant interaction, following on from the pioneering work of Stern (Stern, 1995) and Emde (e.g., Emde and Spicer, 2000). Beebe’s frame-by-frame time-series analysis of mother-infant

interactions at 4-months revealed the role of the intrapersonal over the interpersonal factors in determining attachment security (Beebe et al., 2010b, Beebe et al., 2012). It turned out that mothers’ contingent behaviour with infants’ prior behaviour (contingent responding) was least predictive of attachment security. More predictive were the ways that mothers regulated their own behaviours, how infants coordinated their behaviours with mothers, and the infants regulation of their own behaviours. Infants later categorised as resistant had mothers who impinged on them but these infants also ‘dodged’ as the mothers ‘chased’. Insecure infants showed both higher self-contingency in touch, and lower in engagement. Higher engagement coordination with mother’s engagement characterised insecure (especially resistant infants) but they also manifested lowered engagement coordination with maternal intrusive touch. Disorganized infants showed higher self-contingency in facial affect, but lower in gaze and spatial orientation. Thus, higher self- or interactive contingency is not necessarily better – both lower, or more withdrawn, interactivity and heightened, or vigilant, contingent interaction were associated with insecure dyads. Patterns are complex and contextually moderated. None of this should surprise us. What is surprising is that we ever expected it to be more straightforward. In the same study, a qualitative analysis, which involved identifying ‘behavioral extremes,’ complemented the quantitative (contingency) analysis and turned out to be critical in revealing important maternal contributors to the process of forming insecure attachment. For example, the mother’s exaggerated surprise or smiling response to infant distress predicted disorganized attachment. A characteristic precursor pattern for attachment disorganisation shows the mother responding to the downward turned mouth of the infant by looming and smiling. When the infant responds to looming by breaking eye contact the mother looms and smiles more. The infant now turns its head down, but the mother is still looming and smiling. Finally, when the infant manifests distress, the mother shows surprise. What Beebe’s work, along with considerable other empirical observation, highlights is the mutual construction of attachment relationships by the infant and the carer. This does not assume complete symmetry between mother and child in driving the character of attachment – the mother clearly has ‘a greater
range, control, and flexibility of behaviour than the infant’ (Beebe et al., 2010a, p.113), and particular extremes of maternal behaviour (such as a mother smiling in response to the infant’s distress), may have a disproportionately significant impact in determining attachment status over more frequently occurring but less jarring interactions. Attachment is a fluid, co-constructed and flexible system which is distorted and compromised through static formulations. This is of course close to psychoanalytic formulations.

Work such as Beebe’s can move the “psychoanalytic infant” on from being a hypothetical creature based on the speculative reconstruction of adult narratives, to a picture that is constrained and moderated by actual systematic observations of children yet retains its dynamic character. But the picture is far from straightforward and coherent models shared between psychoanalysis and attachment theory will require the thoughtful integration of qualitative and quantitative observational data. A good example of a successful theoretical elaboration of this kind is Tronick’s (2007) attempt to shift the notion of the IWM into the more subtle and dynamic frame of a mutual regulation model (MRM), which focuses on the moment by moment meaning-making by the dyad; meaning that is constructed through infant-adult interactions which can be highly elusive and ephemeral, non-verbal micro-regulatory processes.

To summarise what we have said so far, as the value of attachment thinking has become well-established, empirically, clinically and theoretically (during the Kuhnian stage of normal science), the future of attachment thinking and research seems to demand a greater refinement of some of the established thinking. And some of the changes that seem to be required of attachment work, which point to less blunt forms of assessment and measurement of attachment and its construction, and one that accommodates developmental fluctuations across the life course, speak to some of psychoanalysis’s concerns about attachment – although an emphasis on the importance of an empirical, evidence-based approach to such developmental work may continue to vex many psychoanalysts.
The challenge of psychotherapy outcome

Meanwhile, one of the questions that has long dogged psychoanalysis, and may well have some bearing on its long-term future as a pursuit that goes beyond the esoteric, is raised by the so-called ‘Do-do’ verdict – the finding that all bone fide psychodynamic interventions are all roughly as effective as one another, no matter how wildly their stated method or underlying theory might differ, or even contradict one another (e.g., Mansell, 2011, Budd and Hughes, 2009). On a practical level, the psychoanalytic world has accommodated the reality of the ‘Do-do’ verdict via the much-vaunted pluralism and spirit of non-denominational openness that has replaced the factional warfare that dominated the field during the middle decades of the twentieth-century. This pluralism is to be applauded, but it is also a symptom of the fact that when it comes to psychoanalysis, no one theory or school of thought can, with any authority, claim to know what works, or why.

Psychoanalysis has retained its identity through a conceptual Pascalian Wager in terms of its explanatory mechanisms. We suggest that developing a stronger understanding about its own processes is the issue that psychoanalysis now confronts – if it is to develop the kind of self-awareness that classical analysis itself so forensically sought, or to redirect Freud’s words to the question of reflecting on what makes psychoanalysis work rather than taking it as a leap of faith: ‘ a consciousness of which one knows nothing seems to me a good deal more absurd than something mental that is unconscious.’

Mentalization theory and the capacity to learn from experience

Mentalizing is the capacity to understand ourselves and others in terms of intentional mental states. It involves an awareness of mental states in oneself or in other people, particularly when it comes to explaining behaviour. That mental states influence behaviour is beyond question. Beliefs, wishes, feelings and thoughts, whether inside or outside our awareness, always determine what we do. Mentalizing is often simplistically understood as synonymous with the capacity for empathy towards other people. In fact, mentalizing involves a spectrum of capacities: it is underpinned by four polarities involving relatively distinct neural circuits: (a) automatic-controlled,
(b) internally-externally-based, (c) mentalizing with regard to self and others and (d) cognitive versus affective (Fonagy and Luyten, 2009). As such, mentalizing provides a comprehensive map of social cognition.

As the human mind was required to respond to ever more challenging, complex and competitive conditions, the exact nature and content of social knowledge cannot be fixed by genetics or constitution; it must be optimized through a prolonged period of development by a group of kin, that is, attachment figures. Mentalization exists not only to permit superior adaptation to the physical environment through facilitating social collaboration and well-functioning kinship groups, but also to support competition for survival when social groups are at odds. Evolution has charged attachment relationships with ensuring the full development of the social brain. The capacity for mentalizing, along with many other social-cognitive capacities, evolves out of the experience of social interaction with caregivers. Increased sophistication in social cognition evolved hand in hand with apparently unrelated aspects of development, such as increased helplessness in infancy, a prolongation of childhood, and the emergence of intensive parenting.

Mentalizing theory was born of attachment theory: attachment contexts provide the ideal conditions – in secure attachment relationships, where attachment figures are interested in the child’s mind – for fostering mentalizing (Fonagy et al., 1991). Thus mentalizing is clearly indebted to attachment theory, and it has contributed to work on attachment by explaining the processes of intergenerational transmission and why secure attachment may be so important in explaining resilience in the face of adversity (Fonagy et al., 1994) and how attachment trauma interfaces with personality disorder (Fonagy, 1998). We have suggested that the carer’s capacity to mentalize the child bolsters the child’s opportunity to develop secure attachment. Against the background of that attachment relationship, the infant is free to explore other subjectivities, including that of his caregiver. Finding himself accurately represented as a thinking and feeling intentional being will in turn ensure that his own capacities for mentalizing will develop well (Fonagy et al., 2002). Recent
elaborations of thinking on mentalizing, however, have taken the argument one step further, and point to yet another important function of attachment relationships: that of the development of epistemic trust and the enabling of social learning in an ever changing environmental context (Fonagy et al., 2014, Fonagy and Luyten, in press).

Fundamental to this new development in mentalization theory is an evolutionarily informed approach to a uniquely human dilemma regarding the transmission of culture and social learning (Heyes and Frith, 2014). As humans developed ever increasing levels of social complexity, the survival of the individual and the group required the transmission of far more demanding levels of social knowledge (Wilson, 1976). Human infants were increasingly born into a world heavily populated with man-made tools whose functional properties, appropriate manner of application or method of reproduction were epistemically opaque; this meant that we came to rely on communication from a trusted source to learn to navigate the use of such objects. This learnability problem is made more difficult by the fact that we have an opposing need for epistemic vigilance – for caution and discrimination on the part of the juvenile-observational learner to prevent them from being tricked or misinformed, intentionally or not (Sperber et al., 2010b). Gergely and Csibra have suggested that it was to overcome this natural epistemic vigilance that we have evolved a human-specific, cue-driven form of social cognitive adaptation dedicated to ensure the efficient transfer of knowledge – this is the theory of natural pedagogy (ToNP) (Csibra and Gergely, 2009, Csibra and Gergely, 2011b). ToNP posits that humans are evolved to both teach and learn new and relevant cultural information. Communication is specifically adapted to allow the transmission of cognitively opaque information – that is, generic knowledge that is robust to interference, is kind generalizable and becomes experienced as shared in the sense that it immediately generates an expectation that others belonging to the social group also possess this knowledge.

We have integrated the ideas of Sperber, Gergely and Csibra with our previous thinking about attachment as a key context for the acquisition of mentalizing skills. We have consistently
suggested that the evolutionary function of the dyadic relationship between parent and human infant goes way beyond ensuring the safety of the latter, as the attachment system was “captured” by evolution to provide a platform for the transmission of cultural knowledge across generations, most particularly understanding about the nature of subjectivity and the symbolic functioning of the human mind (Fonagy et al., 2007, Fonagy et al., 2002). In our view, the major evolutionary advantage of attachment in humans is the opportunity it gives the infant to develop social understanding. Alan Sroufe and Myron Hofer (Hofer, 2003) played a seminal role in extending the scope of attachment theory from an account of the developmental emergence of a set of social expectations to a far broader conception of attachment as an organizer of physiological and brain regulation. Our work simply extends their ideas. Attachment ensures that the brain processes that serve social cognition are appropriately organized and prepared to enable us to live and work with other people. Our theories were partly developed in response to the limits of attachment theory, and see the evolutionary role of attachment as going far beyond the scope which Bowlby specified for this process.

Developing the work of Bertrand Russell (1940) and Sperber and Wilson (1995), Csibra and Gergely propose that the transmission of knowledge is triggered in humans through the use of ostensive cues. Human infants display a species-specific sensitivity in relation to certain nonverbal ostensive behavioral signals (Csibra and Gergely, 2006, Csibra and Gergely, 2009, Csibra and Gergely, 2011a). These are stimuli on the part of the ‘teacher’ that indicates they are communicating relevant cultural information. Infants attend preferentially to such signals and the impact of these signals on their behaviour is readily apparent. Ostensive cues include eye contact, turn-taking contingent reactivity, and the use of a special vocal tone (‘motherese’), all of which appear to trigger a learning state of mind in the infant, a pedagogic stance opening a channel for receiving new information about the social and personally relevant world; they generate an understanding that this information goes beyond specific experiences and that it is generically applicable. There is considerable accumulating evidence to support the view that establishing personal contact with a
young child serves to increase the likelihood that a subsequent communication will be considered to contain personally relevant and generalizable information (Egyed et al., 2013, Kiraly et al., 2013). It is hardly surprising that human infants (as do human adults) learn and internalize knowledge more readily if they experience the teacher as concerned about them, as having their mind in mind.

Intriguing implications follow, for both the theory of psychopathology and the practice of psychodynamic interventions. Ostensive cues create an opportunity for learning because they generate a sense of epistemic trust. Ostensive cues are powerful indicators to young persons that another person has specific interest in them. This interest is reflected above all by contingent responding. Attachment is established through the same process. Thus we have two vital biological mechanisms which may turn out to be closely linked: a mechanism for relating, which is part of emotion regulation, and a mechanism for learning and exploration. The latter Bowlby originally regarded as separate although conditional on the acquisition of an experience of a background of safety (Bowlby, 1988). The contingent responding by the caregiver establishes a specific and irreplaceable bond to the attachment figure while simultaneously establishing a sense of trust in that person as a source of knowledge. A child who has experienced secure attachment also experiences their primary attachment figure as a dependable source of knowledge – as someone who is both truthful and authoritative in their understanding of the surrounding social world. By extension the expectation of being generally responded to sensitively (secure attachment) also creates a general readiness to learn. Not surprising then that secure children learn faster.

Attachment is a special condition for generating epistemic trust; the likelihood that a piece of knowledge we receive will be taken on board as relevant and internalized to modify our representation of the world. Children are eager to learn about the opaque mental world from those around them, but they are prepared to learn most readily about minds in conditions of epistemic trust.
Building on pioneering work by Dan Sperber (Sperber et al., 2010a, Wilson and Sperber, 2012), we consider the interpersonal underpinnings of efficient learning from experience, at least in part, to be based around deferential authority assigned to certain individuals in whom we place epistemic trust. Secure attachment experiences are based on having been mentalized or felt understood. Mentalizing or acting contingently in accordance with the internal state of another person is an ostensive cue and therefore may be expected to generate and also foster epistemic trust – that is, an individual’s disposition to consider new knowledge from another person as trustworthy, generalizable, and relevant to the self. In other words, natural selection may have hit upon attachment to mediate the reliable identification of individuals to be trusted with transmission of precious cultural knowledge (‘memes’) (Wilson, 1976, Wilson and Wilson, 2007) from one generation to the next. Secure attachment helps to create a benign condition for the relaxation of epistemic vigilance because it reflects sensitive and appropriate ostensive cueing.

While attachment may be a key mechanism for mediating epistemic trust, it is secondary to an underlying biological process preserved by natural selection. Secure attachment is not a necessary condition for generating epistemic trust but it may be sufficient to do so, and, further, it is the most pervasive mechanism in early childhood because it is a highly evolutionarily effective indicator of trustworthiness. Therefore, what has been underemphasized within attachment theory, in our opinion, is that in a securely attached child, the caregiver is understood to be a reliable source of knowledge. The consistent emotional responses of a sensitive caregiver are clearly enacted to the child via the caregiver’s ostensive cues – often in the form of marked mirroring (the exaggerated facial displays and the use of the voice which reflect how the infant is feeling back to the infant, but in a ‘play-acting’ manner). By marked mirroring, the adult is describing the child’s emotional state as it is being felt by the child, in a manner which simultaneously makes it clear that it is the child’s feelings that are being enacted and understood, not the adult’s. It is via these interactions that the child learns the content of mental states – to understand their own minds and to begin to imagine what might be happening in other people’s minds. In relation to the issue of the neglect of sexuality
and aggression in attachment thinking, mentalizing has provided a way of integrating this with traditional psychoanalysis, as it is emphasized that sexuality and aggression are the least marked mirrored in development, thus always remaining somewhat alien, and so prone to acting out. (Fonagy, 2008) (Fonagy and Luyten, in press)

According to this evolutionary perspective, a particular attachment style could be seen less as a measurement of the extent to which the parent was able to succeed in generating infant attachment security. The attachment style adopted by the child should, more broadly, be seen as an adaptive learning outcome concerning what might be the most appropriate method for social survival in a complex interpersonal world (Simpson and Belsky, 2008, Belsky, 2006b, Mikulincer and Shaver, 2007). It seems conceivable that an avoidant/dismissive model of attachment might be more protective in certain environments than a secure one. Similarly, the anxious/preoccupied style may be an effective means of ensuring a child learns to effectively harness interpersonal attention and resources in a context of resource uncertainty. Taking the argument further, we can postulate that even serious personality disorders such as BPD, while conspicuously dysfunctional in our normative social setting, may have adaptive benefits for individuals living in an emergency milieu characterized by high levels of interpersonal violence, where there is a need for extreme vigilance on issues of self-protection and where there is significant benefit in being able to form intense emotional relationships, which might elicit critical protection or resource supply very quickly. The mentalizing strengths that have been noted in many individuals with BPD – a tendency to be able to make quick inferences of other people’s mental states on the basis of their immediate visual and emotional cues, a hypersensitivity to facial expressions, hyper-reactivity to positive and emotional stimuli – are all suggestive of a mentalizing profile that may be an adaptation to functioning in a threatening or high-risk environment.
Ways forward

The clinical implications of this thinking on mentalizing, attachment and epistemic trusts fall into two categories. The first involves the conceptualization of psychopathology, the ways in which we structure and understand different forms of mental health disorder, their comorbidity, and the nature of psychopathology across the life course. The second clinical implication arising from our proposed theory concerns the organization and delivery of clinical interventions; what makes therapy effective and what are the surrounding conditions necessary to support an effective outcome.

The conceptualization of psychopathology

A striking challenge for attachment researchers is that little evidence has been found showing a relationship between different attachment styles and particular forms of psychopathology. In a meta-analysis, Fearon et al. found a significant association between insecure attachment and later externalizing problems in children; with a larger effect size for boys and for disorganized children (Fearon et al., 2010). Beyond this, no specific relationship has as yet been found between attachment styles and specific psychopathologies. A broader challenge for our thinking about psychopathology arises from the fact that when we consider an individual’s psychiatric history over their life-course, it rarely follows the discrete, symptom-led and time-limited categories that extant research uses when conceptualizing specific disorder. Instead – particularly in cases of more severe impairment – an individual’s clinical presentation changes over time, one typical example being a progression from adolescent conduct disorder to adult depression.

This lack of specificity may relate to compelling evidence presented by Caspi and colleagues suggesting that there is, in fact, one general psychopathology factor in the structure of psychiatric disorders (Caspi et al., 2013). In their longitudinal study based in Dunedin, New Zealand, Caspi examined the structure of psychopathology from adolescence to mid-life, examining dimensionality, persistence, co-occurrence and sequential comorbidity. The study found that psychiatric disorders
were more convincingly explained by one General Psychopathology factor (labelled the p factor as a conceptual parallel to the g factor, the well-established dimension by which general intelligence is understood). A higher p factor score is associated with increased severity of impairment: ‘Higher scores on this dimension were associated with more life impairment, greater familiality, worse developmental histories, and more compromised early-life brain function’ (Caspi et al., 2013, p.131). The p factor concept convincingly explains why discovering isolated causes, consequences or biomarkers and specific, tailored treatments for psychiatric disorders has proved so elusive for the field.

The p factor is a statistical construct. The question that follows from this initial conceptualization is: what is the p factor, what does it measure? According to our latest thinking, we propose that the p factor is a measurement of epistemic trust: an individual with a high p factor score is one in a state of epistemic hypervigilance and epistemic mistrust. A depressed patient with a low p factor may recover with the help of CBT delivered via an e-platform – they are relatively easy to reach in terms of treatment because they are open to social learning in the form of therapeutic intervention. However, a depressed patient with a high p factor – suffering from high levels of comorbidity, longer-term difficulties and greater impairment – will demonstrate treatment resistance because of their high levels of epistemic mistrust, or outright epistemic freezing. We consider it likely that such patients require more long-term therapy that is highly mentalizing and rich in ostensive cues, which will serve to stimulate epistemic trust and openness.

We posit that the developmental relationship between epistemic hypervigilance and the nature of psychopathology most typically occurs when social adversity – most deeply through trauma or maltreatment – destroys the individual’s trust in incoming social knowledge of all kinds. The individual becomes rigid and hard-to-reach on an interpersonal level because they cannot accept new information as relevant to them across social contexts, whether this new information is received via their own experiences, communication from attachment figures, or anyone else.
According to this model, we conceptualize personality disorder as a form of inaccessibility to cultural communication (Fonagy et al., 2014) (Fonagy and Luyten, in press). The epistemic mistrust which can follow experiences of neglect, maltreatment or abuse can lead to an epistemic hunger as well as mistrust, with profound clinical implications that the therapist ignores at their peril – potentially leading to the iatrogenic effects that can arise from the activation of a disorganized attachment system, for example in the treatment of borderline personality disorder (BPD) (Fonagy and Luyten, in press). The rigidity and unreachable quality so often found in cases of BPD, which has historically made this patient group so hard to help (often causing intense feelings of frustration in therapists), is caused by an inability to trust what they hear.

Clinical Implications

The theory of epistemic trust as the underlying structure of psychopathology has important implications for clinical interventions, and the role of attachment and mentalizing within them. Using BPD as an example, we suggest that an increase in epistemic trust is the driving force of therapeutic change. However, we also argue that reopening epistemic trust in a therapeutic context is merely one part of a process of social learning, or rather social relearning, that needs to be supported in the patient’s wider social environment for there to be any chance of sustained or meaningful change. This speaks to the dilemma of the ‘Do-do’ verdict, discussed above. In proposing that epistemic mistrust might constitute the p factor that underlies psychopathology, we also consider that the reopening of epistemic trust may be at the heart of all effective therapeutic interventions. We further propose that this is a three-stage process of change, in which attachment, mentalizing and the social environment synergistically interact with the re-emergence of epistemic trust.

*Communication System 1: The communication of content*

The first stage of any effective intervention involves the transmission of substantive content to the patient about the nature of their state in a way that is coherent and credible enough for the patient
to accept. This stage can involve different explanatory models – for example, early schemas, invalidating experiences, object relations, and current attachment experiences – but the point is that the model conveyed to the patient is experienced by them as personally relevant. The patient has the experience of their agentive self being understood through a therapeutic form of marked mirroring. As well as providing a form of relevant social communication, the very process of applying and transmitting this knowledge involves a subtle and rich process of ostensive cueing on the part of the therapist. It requires the therapist to be able to mentalize the patient effectively, to be able to understand the patient’s state well enough to be able to respond with the right forms of ostensive cues and to formulate the model – the social knowledge of the intervention – in a manner that fits with the patient’s own experiences of their own mental state. The actual content of this form of social knowledge can provide extremely valuable ways for the patient to understand – to mentalize – themselves and their reactions to others. But the generic value of communicating this social knowledge, which falls across different models of treatment, lies in the patient’s experience of being recognised. This is what skilful, sensitive therapeutic mentalizing involves, and it results in a relaxation of epistemic mistrust.

**Communication System 2: The re-emergence of robust mentalizing**

In communicating social knowledge in the form of the therapeutic model, the therapist is intrinsically working to recognize the patient’s agentive self; is marking the patient’s experiences by presenting the information in a way that acknowledges the patient’s emotional state; and, is using ostensive cues to convey to the patient both the personal relevance and the wider social value of the therapeutic model. The therapist is in effect helping the patient to reopen their evolutionarily-endowed ability to receive social communication from attachment figures, a capacity which has been lost for reasons of environmental adversity, genetic propensity or both. By effectively mentalizing the patient, the therapist is modelling how they mentalize in an open, trustworthy and relatively low-arousal environment. This creates a virtuous cycle in which the patient, in being sensitively responded to, retreats from their epistemic isolation and may begin to exercise and
develop their own mentalizing skills. The process summarized here is simplifying what is often a complex, non-linear progression, but it is a process in which, ideally, the patient experiences an epistemic shift and develops more balanced and robust mentalizing capacities.

However, we would like to suggest (in a way that may at first appear counter-intuitive for avowed supporters of mentalizing-based therapy) that this improvement in mentalizing is not in itself the clinching objective of enduringly effective therapy. The significance of improved mentalizing is that it enables the patient to learn from their wider social world, which takes us to the next stage.

**Communication System 3: The re-emergence of social learning beyond therapy**

The improved mentalizing that results from a successful intervention, ideally, brings about improved social relations and experiences outside the consulting room. Improved epistemic trust and the breaking down of the rigidity with which social experiences are interpreted and responded to, paves the way for the patient to accumulate experiences of social interaction that are benign, or that are at least manageable in terms of maintaining resilient mentalizing. This creates a virtuous circle of growing robustness in mentalizing and improving mentalizing capacity. This final, critical stage of social learning beyond therapy is highly contingent on the individual’s social environment being benign, or ‘benign enough’. Therapeutic change can only be sustained, according to this thinking, if the patient is able to use, and even to change (through the seeking out of more mentalizing relationships) their social environment in a way that allows them to continue to relax epistemic hypervigilance and foster their mentalizing strengths.

**Conclusion**

Fifteen years ago, in *Attachment Theory and Psychoanalysis*, the approach to the bad blood that tainted the relationship between the two fields was to seek out areas of conceptual rapprochement and conciliation. Now – and not entirely through attrition – we seek to locate the debate not in terms of bad blood but of the need for new blood. Both attachment and psychoanalysis must, if
either or both field are to retain their intellectual vigour and relevance, rethink their approach to psychopathology in a manner that moves us on from a descriptive, category-driven approach, the legacy of an essentially nineteenth-century medical mind-set. Conceptualizing the infant’s mind as a highly contingent and active, evolutionary-driven tool for the pursuit of social knowledge and relationships, as well as for the discharge of drives, allows us not only to understand clinical psychopathology, but also the psychopathology of modern, everyday life.


KUHNEN, P. in press. Unholy questions about five central tenets of psychoanalysis that need to be empirically verified. *Psychoanalytic Inquiry*.


