

November 1998

*The value of this manual is likely to depend upon accompanying training by the authors,
or by others whom they have trained and found to be reliable*

REFLECTIVE-FUNCTIONING MANUAL

Version 5

FOR APPLICATION TO ADULT ATTACHMENT INTERVIEWS

Peter Fonagy, Mary Target, Howard Steele and Miriam Steele

Address for correspondence:-

Professor Peter Fonagy
Psychoanalysis Unit
Sub-Department of Clinical Health Psychology
University College London
Gower Street, London, WC1E 6BT

Table of Contents

1	Introduction	4
1.1	Defining the term	4
1.2	The historical roots of the concept	4
1.3	Metacognition, mentalisation, and reflective-functioning	6
1.3.1	Mentalisation and the self.....	7
1.3.2	The development of mentalising skills	7
1.4	Mentalising skills and self reflection	9
2	Why is mentalisation or reflective-functioning so important?	10
2.1	Behaviour becomes predictable.....	10
2.2	Reflective-functioning promotes and maintains attachment security.....	10
2.3	Reflective-functioning facilitates the appearance-reality distinction.....	10
2.4	Reflective-functioning enhances communication.....	11
2.5	Reflective-functioning encourages meaningful connections between the internal and external worlds	11
3	Validation of the measure.....	11
3.1	The London Parent-Child Project.....	11
3.2	The Cassel Hospital Study	16
3.3	The Prison Health Care Centre Study	18
3.4	The use of RF ratings in predicting the child's capacity to mentalize	18
4	Illustrations of moderate to high reflective-function	19
4.1	Awareness of the nature of mental states.....	19
4.1.1	The opaqueness of mental states	20
4.1.2	Mental states as susceptible to disguise.....	20
4.1.3	Recognition of the limitations on insight.....	20
4.1.4	Mental states tied to expressions of appropriate normative judgements.....	20
4.1.5	Awareness of the defensive nature of certain mental states	21
4.2	The explicit effort to tease out mental states underlying behaviour	21
4.2.1	Accurate attributions of mental states to others.....	21
4.2.2	Envisioning the possibility that feelings concerning a situation may be unrelated to observable aspects of it	21
4.2.3	Recognition of diverse perspectives.....	21
4.2.4	Taking into account one's own mental state in interpreting others' behaviour.....	22
4.2.5	Evaluating mental states from point of view of its impact on behaviour of self and/or other	22
4.2.6	Taking into account how others perceive oneself	22
4.2.7	A freshness of recall and thinking about mental states	22
4.3	Recognising developmental aspects of mental states	23
4.3.1	Taking an intergenerational perspective, making links across generations	23

4.3.2	Taking a developmental perspective	23
4.3.3	Revising thoughts and feelings about childhood in light of understanding gained since childhood	23
4.3.4	Envisioning changes of mental states between past and present, and present and future.....	24
4.3.5	Envisioning transactional processes between parent and child	24
4.3.6	Understanding factors which developmentally determine affect regulation	24
4.3.7	Awareness of family dynamics	24
4.4	Mental states in relation to the interviewer	25
4.4.1	Acknowledging the separateness of minds	25
4.4.2	Not assuming knowledge	25
4.4.3	Emotional attunement	26
5	General considerations	26
5.1	Only explicitly reflective statements qualify for high ratings.....	26
5.2	Learned, rote, or clichéd statements do not qualify for high ratings.....	27
5.3	Reference to personality or a relationship, in the absence of specific reference to mental states, does not qualify for a high rating	27
5.4	Avoiding thinking for the subject.....	27
5.5	Diagnoses should not be accepted as a shorthand for mental states	28
6	Illustrations of negative or limited reflective-functioning.....	28
6.1	Possible implications of sub-types of impoverished reflective-functioning.....	28
6.2	Rejection of RF	29
6.3	Unintegrated, bizarre or inappropriate RF	30
6.4	Disavowal of RF	31
6.5	Distorting or self-serving RF	32
6.6	Naive or simplistic RF.....	34
6.7	Overly-analytical or hyperactive RF	35
7	Rating passages	37
7.1	Rules for identifying passages: Demand vs. permit questions.....	37
7.2	Guidelines for rating identified passages	38
7.3	Rules for aggregating reflective-functioning ratings into a single score for each Adult Attachment Interview	42
7.4	The overall interview rating	43
	Acknowledgements.....	48
	References.....	49
	Appendix 1: Outline of scoring procedures.....	50

1 INTRODUCTION

1.1 Defining the term

The term reflective function (RF) refers to the psychological processes underlying the capacity to mentalize, a concept which has been described in both the psychoanalytic (Fonagy, 1991; Fonagy & Higgitt, 1989) and cognitive psychology literatures (Morton & Frith, 1995). Reflective functioning or mentalization is the active expression of this psychological capacity intimately related to the representation of the self (Fonagy & Target, 1995, 1996; Target & Fonagy, 1996). RF involves both a self-reflective and an interpersonal component that ideally provides the individual with a well-developed capacity to distinguish inner from outer reality, pretend from ‘real’ modes of functioning, intra-personal mental and emotional processes from interpersonal communications. Because of the inherently interpersonal origins to how the reflective capacity develops and expresses itself, this manual refers to reflective functioning, and no longer of reflective-self functioning (see Fonagy, Steele, Moran, Steele, & Higgitt, 1991b) as the latter term is too easily reduced to self-reflection which is only part of what is intended by the concept.

1.2 The historical roots of the concept

Various notions have been introduced in the psychoanalytic literature to denote mental processes that overlap with the construct of mentalisation, underpinned by RF. While space does not permit an exhaustive review some of these should be mentioned as these may be helpful to the reader attempting to make links between the current constructs and those proposed by other writers. All such notions derive from Freud’s initial concept of “Bindung” or linking. In Freud’s distinction of primary and secondary processes (Freud, 1911), he stressed that “Bindung” was both a qualitative change from a physical (immediate) to a psychical associative quality of linking and that the psychical working out or representing of internal state of affairs (conceived of in energetic terms) failed in various ways (Freud, 1914). Some might argue that Melanie Klein’s notion of the depressive position (Klein, 1945) is at least analogous to the notion of the acquisition of RF which necessarily entails the recognition of hurt and suffering in the other as well as that of one’s own role in the process. Wilfred Bion (1962a,b) in describing the “alpha-function” delineated the transformation of internal events experienced as concrete (“beta-elements”) into tolerable thinkable experiences. Similarly to the current conception Bion also saw the mother-child relationship as at the root of the symbolic capacity. Winnicott (1962) also recognised the importance of the care-giver’s *psychological* understanding of the infant for the emergence of the true self. Winnicott was also foremost amongst psychoanalytic theorists of self development (e.g. Fairbairn, 1952; Kohut, 1977) in recognising that the psychological self develops through perception of oneself in another person’s mind as thinking and feeling. Parents who cannot reflect with understanding on their children’s inner experiences, and respond accordingly, deprive their children of a core psychological structure which they need to build a viable sense of self.

Independently, French psychoanalysts developed a notion of mentalisation largely formulated

from the economic point of view. Pierre Marty discussed mentalisation as a protective buffer in the preconscious system which prevents progressive disorganisation (Marty, 1968). He considers mentalisation as connecting drive excitations and mental representations and thereby creating both “fluidity” and “constancy” (Marty, 1990, 1991). Mentalisation ensures freedom in the use of associations as well as permanence and stability. At the same time Pierre Luquet (1981, 1988) discussed the development of different forms of thinking and the reorganisation of inner experience alongside this development. In his chapter on a theory of language (Luquet, 1987) he distinguished primary mentalisation (which we would consider the absence of RF) with secondary symbolic mentalisation. While this form of mentalisation was still seen as closely connected to sensory data and primary unconscious fantasies it was nevertheless also seen as representative of these processes and observable in dreams, art and play. His third level was verbal thought which he considered most distant from bodily processes. Similar ideas were proposed by André Green (1975), Hanna Segal (1957) and Joyce McDougall (1978), and more recently by Frosch (1995), Busch (1995) **Fehler! Verweisquelle konnte nicht gefunden werden.** and Auerbach (1993; 1996).

Developmentalists over the past ten years have drawn our attention to the universal and remarkable capacity of young children interpret the behaviour of all agents (themselves as well as others) in terms of putative mental states. RF is the developmental acquisition that permits children to respond not only to another person’s behaviour, but to the children’s conception of others’ beliefs, feelings, attitudes, desires, hopes, knowledge, imagination, pretence, deceit, intentions, plans and so on. RF or mentalization enables children to “read” other people’s minds (Baron-Cohen, 1995; Baron-Cohen, Tager-Flusberg, & Cohen, 1993; Morton & Frith, 1995). By attributing mental states to others, children make people’s behaviour meaningful and predictable. As children learn to understand other people’s behaviour, they can flexibly activate, from the multiple sets of self-other representations they have organised on the basis of prior experience, the one(s) best suited to respond adaptively to particular interpersonal transactions.

The notion of RF is rooted in Dennett’s (1987; 1988; 1978) proposal that three stances are available in the prediction of behaviour: the physical stance, the design stance and the intentional stance. He takes predicting the behaviour of a chess-playing computer as his example. At its simplest this can be based on knowledge of the physical properties of the machine (the physical stance). The design stance would be based on knowledge of the design of the computer including the programming which went into its development. The third approach consists of predicting what might be the computer’s most rational move. Here we attribute to the computer certain beliefs and desires; in other words regulation by intentional states. Dennett’s thesis is that explanation in terms of such states of meaning provides good grounds for predicting human behaviour -- the only grounds accessible to all of us – this knowledge is embodied in the theory of mind of folk psychology (Churchland, 1986; Fodor, 1987; Mele, 1992).

Theory of mind is an interconnected set of beliefs and desires, attributed to explain a person’s behaviour. The theory of mind concept has great generalisability and explanatory value. Recent philosophers of mind (Hopkins, 1992; Wollheim, 1995) have extended Dennett’s approach to

unconscious processes. They illustrated that one of Freud's substantive contributions was to extend folk psychology to unconscious mental states, a theory of unconscious mind, thus making those aspects of behaviour meaningful which - using the ordinary constructs of intentionality - make little sense (e.g. dreams, neurotic symptoms, humour). These behaviours may be understood if we add unconscious beliefs, thoughts, feelings and desires to our everyday model of the mind.

Extending these ideas, we consider RF to be the mental function which organises the experience of one's own and others' behaviour in terms of mental state constructs. RF concerns knowledge of the nature of experiences which give rise to certain beliefs and emotions, of emotions and beliefs consequent upon particular experiences, of likely behaviours given knowledge of beliefs and desires, of the expectable transactional relationships between beliefs and emotions, and of feelings and beliefs characteristic of particular developmental phases or relationships. Its essence is not that the individual should be able to articulate this theoretically, and this is clear in our operationalisation below. It is our view that individuals differ in the extent to which they are able to go beyond immediately known phenomena to give an account of their own or others' actions in terms of beliefs, desires, plans and so on. This, undoubtedly high level, cognitive capacity is an important determinant of individual differences in self-organisation as it is intimately involved with many defining features of selfhood such as self consciousness, autonomy, freedom and responsibility (Bolton & Hill, 1996; Cassam, 1994). Intentional stance, in the broad sense considered here (i.e. including apparently irrational unconscious acts), explains intentional behaviour (Davidson, 1980) in oneself and therefore creates the continuity of self experience which is the underpinning of a coherent self structure.

Much experimental work has been performed which demonstrates the theory of mind is acquired during the fourth year of life (e.g. Baron-Cohen et al., 1993; Morton & Frith, 1995). Our formulation differs from most developmentalists in considering RF to be a developmental achievement which is never fully acquired i.e. maintained across situations. Also our emphasis along with other psychoanalytic authors is on the self organising quality of mentalising and the implications which individual differences in this capacity have for our understanding of psychological disorder.

1.3 Metacognition, mentalisation, and reflective-functioning

The rating scale evolved in part from Mary Main's (1991) seminal chapter on "metacognitive monitoring and singular versus multiple models of attachment." Main's concept principally refers to the activity of self-monitoring in the context of the Adult Attachment Interview (Main & Goldwyn, 1995) protocol. The reflective-functioning scale could be conceived of as providing operationalised definitions of individual differences in adults' metacognitive capacities. Metacognition, mentalisation and reflective-functioning are seen as expressions of the RF on which, in large part, depends the development of the self who thinks and feels.

1.3.1 Mentalisation and the self

Mentalising refers to the capacity to perceive and understand oneself and others in terms of

mental states (feelings, beliefs, intentions and desires). It also refers to the capacity to reason about one's own and others' behaviour in terms of mental states, i.e. reflection. Mental processes during the first years of life are largely pre-reflective; we nonetheless regard the reflective capacity -- which evolves during this same period -- as belonging to the core self structure. The emergence and full development of the RF depends upon the caregiver's capacity to more-or-less accurately perceive intentionality in the infant. The capacity for reflection is seen as influencing the quality of psychic reality experienced by the individual, and accounts for the richness and diversity of inner experience.

1.3.2 The development of mentalising skills

Mentalisation normally comes about through the child's experience of his mental states being reflected on, prototypically through experience of affect-laden interactions with the caregiver. These range in intensity and complexity depending on the age of the child. The most important interpersonal experiences with caregivers, which ordinarily confirm the child's sense of being reflected in the mind and behaviour of the other, are likely to be those following the child's expression of distress (Gergely & Watson, 1996). For example, in the two month old it is the sensitivity to an infant's cry, whereas in the older child it can occur within the context of a limit-setting response by the parent to one of the child's transgressions. In both cases, the parent reflects upon the child's behaviour, and responds in a way that at once soothes the child's distress (promoting intimacy and sameness) and also suggests a mode of coping (promoting autonomy and separateness). Thus mentalisation by the parent provides or confronts children with a presentation of the contents of the parent's mind that is *both the same and different* from the contents of the child's mind.

Pretend play with a parent provides another avenue for the development of mentalisation, for it enhances the child's integration of inner and outer reality. The child is helped to see these two spheres of existence as linked, yet also different in important ways. This has the consequence of helping the child toward relying less on equating the inner and outer worlds (e.g., significant delusions in older children and adults), or splitting off one world from the other (e.g., isolation or dissociation in older children and adults). In play, the caregiver may give children's ideas and feelings (when they are "only pretending") a link with reality, by showing they can share in the play despite their membership in the grown-up external world. The parent thus shows that reality may be distorted by acting upon it in playful ways. Through this playfulness, real mental experiences are introduced. This is the great appeal and survival value of play and all pursuits in the transitional realm (Winnicott, 1971) as play promotes enhanced 'real' mental experiences. Such experiences are crucial to the individual's attempts to accurately process outer reality and shape or modify the inner landscape (see Emde, Kubicek, & Oppenheim, 1997). Notably, symbolic development generally, and specifically the frequency of usage of mental state terms is radically impoverished in maltreated children (see Cicchetti & Beeghly, 1987).

There is an evident lack of consistency in the use of RF. For example, a psychoanalyst who is highly reflective in a clinical context may be far less so in the context of discussing his own family relationships. A framework for understanding this is provided by 'dynamic skills theory' (Fischer & Farrar, 1987; Fischer, Kenny, & Pipp, 1990) which depicts development as a

person's elaboration of progressively more complex control systems (skills). RF may be readily conceived of as one such control system, critical to the organisation of the self. Within dynamic skills theory, RF would be seen as not simply a property of the person, but of the person and situation together, because all skills are composed of both the person's activities and the situations or context within which these occur. Particular tasks, specific events, other people, as well as culture are seen as part of the skill. Further, the development of a skill is not seen as progression along any singular path, determined by maturation. Rather, RF, as a skill, evolves through varied pathways, moulded by many dynamically interacting influences, such as the individual's emotions, social interaction, family relationships and environment, important social groups, the reactions of the wider social world, etc. (Fischer, Knight, & Van Parys, 1993).

RF is a strand within the developmental web, one of the many distinct control systems that are neither strongly connected with each other, nor co-ordinated or integrated (Fischer & Pipp, 1984). The "fractionation" or splitting of all abilities as a function of tasks and domains is well demonstrated, and we might expect RF to be subject to the same kind of developmental *décalage* (unevenness) which characterises the rest of cognitive development (Flavell, 1982). Fractionation refers to the tendency for a person not to co-ordinate skills or experiences that are naturally separate but may be thought of as belonging together by some external criterion (Fischer & Ayoub, 1994). Just as the understanding of conservation of liquid does not generalise to conservation of area, reflective capacity in one domain of interpersonal interaction should not be expected to generalise to others. RF does not begin as a general capacity, but is a particular skill tied to the task and domain where it is learned, a specific category of relationship. RF as a skill may be more or less present in situations as a function of contextual support and emotional state which push an individual up or down a developmental strand. Differences in the meaning of an interaction as well as its physical context can lead to fractionation. For example, the concept of conservation of liquid may not generalise between the experimental task and one involving helping a thirsty friend, even if both involve pouring a glass of orange juice (Rogoff, 1990). Fractionation does not disappear entirely with development either in general or in the specific case of reflectiveness. It is clearly possible for task-based skills such as RF to come to be co-ordinated but this should not be seen as automatic. Unevenness across situations is likely to remain prevalent even in adults, especially when they are emotional (Fischer & Ayoub, 1994).

Normal development is from fractionation towards integration, which involves the construction of specific co-ordinations amongst previously separate skills and provides the foundation for more complex, sophisticated control systems (Bidell & Fischer, 1994). Abnormalities of RF, the continued use of a non-reflective rather than a mentalizing model for predicting behaviour, should not then be seen as either a consequence of 'arrest and fixation' at an early stage, or a 'regression' to that stage. Deficits in RF may be expected to develop with alternative strategies of increasing complexity over time, in a manner similar to other skills. For instance, the skill for limited reflectiveness developed by the child to anticipate and forestall maltreatment and its painful physical and psychological impact would be adaptive in their particular world but would be expected to produce sophisticated forms of difficulty rather than straightforward adaptations in other contexts (Noam, 1990). The ability to be reflective in general, but to show only minimal reflectiveness in the context of one's own childhood with reference to the mental states of

‘caregiving’ others or in specific relationships which reactivate the same schemata, could be a result of natural fractionation. Unevenness or splitting of reflective ability could also be the consequence of an active (purposeful, conscious or unconscious) attempt on the part of the individual not to co-ordinate or generalise RF to specific relationship domains. Here the unevenness is “a developmental achievement”, in that the person must create a co-ordination in order actively to keep contexts which would naturally move toward integration, separate. The split is context and affect dependent, and skills developed within one context will not be necessarily matched by similar abilities in others. Within an attachment theory framework we might say that the self is organised so that certain internal working models include considerable reflective components – expectations incorporating the mental states of self and other – while other working models of relationships appear impoverished, indicating only minimal mentalizing skills. In the latter contexts the subject will offer only stereotyped, simple, concrete, low level descriptions. This does not imply developmental delay or regression; rather it suggests a remarkably complex ability to co-ordinate two distinct levels of functioning. Thus to talk of deficit or absence of a capacity is undoubtedly an oversimplification. Measures of global abilities may not yield differences between groups. Efforts at going beyond mere clinical impression in terms of measurement and quantification has to take on board the situational and interpersonal specificity of the failure of RF.

1.4 Mentalising skills and self reflection

It is important that RF is not conflated with introspection. Bolton and Hill (1996) note that the weakness of introspection is to define mental states in terms of consciousness or self report rather than, as here, in terms of their capacity to make sense of, and thus regulate, behaviour. Introspection or self reflection is quite different from RF as the latter is an automatic procedure, unconsciously invoked in interpreting human action. We see it as an over-learned skill, which may be systematically misleading in a way much more difficult to detect and correct than mistakes in conscious attributions would be. RF similarly lends a shape and coherence to self-organisation which is entirely outside awareness, in contrast to introspection which has a clear impact on experience of oneself. Procedural knowledge of minds in general, rather than declarative self knowledge, is the defining feature.

2 WHY IS MENTALISATION OR REFLECTIVE-FUNCTIONING SO IMPORTANT?

2.1 Behaviour becomes predictable

Firstly, through the attribution of thoughts and feelings, mentalisation enables the individual to see people’s actions as meaningful. Actions then become predictable, which in turn reduces dependency on others to explicitly detail the meaning of their actions. The child of around four or five is frequently able to understand what the mother is doing and why, without her needing constantly to bear his limited perspective in mind as when a mother conveys the temporary limitations upon her capacity for shared interactions, e.g. “Not now, I’m cooking dinner”. This allows both child and caregiver to attain increasing mental and physical autonomy, needing to

refer far less to each other because children have the mentalising skills to appreciate or borrow their caregivers' understanding. But where does this skill within the child come from? In part, no doubt the answer lies within the maturational framework provided by cognitive developmentalists (e.g. Piaget, Flavell and their followers). These reflections come with development to be internalised and organise the child's subjective experience.

2.2 Reflective-functioning promotes and maintains attachment security

We have empirically demonstrated that individual differences in these skills at appreciating the mental and emotional states of others are rooted in the presence or absence of attachment security. For example, we have demonstrated that high scorers on the RF scale are more likely to have their child securely attached to them than low scorers. We believe that parents' capacity to reflect on the intentionality of their children's behaviour enhances children's self-control and affect regulation (Fonagy, Steele, & Steele, 1991a). Further, we have shown that belief-desire reasoning or mentalising in the child at five years is predicted by both children's and their parents' previously assessed attachment security (Fonagy, 1997a).

2.3 Reflective-functioning facilitates the appearance-reality distinction

Reflection helps the child distinguish between appearance and reality. While this may not be important in all contexts, we believe that in cases of maltreatment or trauma, it allows the child to survive psychologically, and relieves the pressure to enact the experience in concrete ways. Once the child is able to evaluate and use mental representations, he can separate or modify perceptions of, for example, the maltreating parent from perceptions of the self (e.g. "he was unloving but I am not unlovable"). By attributing the origins of ideas and feelings to themselves and/or others, the child makes the human world more explicable. Until children are able to manage this task of reflection, they remain vulnerable to their own and others' immediate emotional reactions. Prior to the development of the RF, inconsistency or hostility from others is more likely to be taken at face value as showing something bad about the child. In contrast, if the child is able to attribute a withdrawn, unhappy mother's apparently rejecting behaviour to her emotional state, rather than to himself as bad and unstimulating, the child may be protected from lasting injury to his view of himself. Indirect support for this view is our finding that mothers with significant experiences of childhood deprivation were unlikely to have insecurely attached children if they were able to reflect accurately or believably upon the interpersonal and intra-personal origins on their experiences (Fonagy, Steele, Steele, Higgitt, & Target, 1994).

2.4 Reflective-functioning enhances communication

Without a clear representation of the mental state of the other, communication must be profoundly limited. The philosopher Grice (1975) formulated the overriding principle of conversation as one of collaboration, whereby the effective speaker needs to bear in mind the point of view of the other person. This perspective has been evocatively elaborated in the writings of Mary Main and her colleagues in the construction of their manual for rating Adult Attachment Interviews (Main & Goldwyn, 1994), particularly with respect to the construct of 'coherence' to which the construct of reflective-functioning is closely linked.

2.5 Reflective-functioning encourages meaningful connections between the internal and external worlds

Finally, mentalising can help an individual to achieve deeper experiences with others, and ultimately a life experienced as more meaningful. We think it is the successful connecting of internal and external which allows belief to be endowed with meaning which is emotionally alive, but manageable. A partial failure to achieve this integration can lead to neurotic states; in more profound and pervasive failures of integration, reality may be experienced as emotionally meaningless, other people and the self are related to as things, and the relating itself occurs at a very concrete level characteristic of severe personality disorders. In the extreme, the individual may be unable to treat themselves or others as motivated by mental states, resulting in a personality organisation sometimes denoted as borderline (Fonagy, 1989, 1991; Fonagy & Higgitt, 1989).

3 VALIDATION OF THE MEASURE

3.1 The London Parent-Child Project

The scale was originally developed as part of the London Parent-Child Project. 100 pregnant women were recruited from the Obstetrics and Gynaecology Department of University College Hospital during the autumn and winter of 1987. Selection criteria included primiparous status, current cohabitation with the child's father, fluency in English and agreement on the part of the father to participate. Approximately 50% of those approached agreed to participate. The sample is described in Fonagy, Steele & Steele (1991a). The sample was mainly middle-class, in the late 20s or early 30s. Subjects were interviewed in their homes, during the pregnancy. Ratings of the interviews with the mothers were performed before the child was one year, and those with the fathers before the child was 18 months. At these points, the baby and mother or father respectively were assessed using the Strange Situation procedure (Ainsworth, Blehar, Waters, & Wall, 1978), to ascertain infant security of attachment.

Each mother interview was rated by four judges, and each father interview by three. The table below shows correlation coefficients between each pair of judges on a nine-point scale. The scale points were defined broadly as in the present manual, with the exception of the allowance for absent (0) or negative (-1) RF, which were added subsequently after rating further groups of interviews from clinical samples.

	Judge A	Judge B	Judge C
Judge B	.835		
Judge C	.600	.590	
Judge D	.665	.600	.749

Table 1. Correlation coefficients between four judges for 100 mother interviews

Each father interview was rated by three judges. The rating manual was improved with stricter definitions and examples added, in-between these two sets of ratings, but the scale remained essentially the same.

	Judge A	Judge B
Judge B	.891	
Judge C	.832	.792

Table 2. Correlation coefficients between three judges for 100 father interviews

We compared RF ratings from the London Parent-Child sample with a number of their demographic characteristics. On the whole, there was a remarkable absence of relationships, with the exceptions of relatively low correlations between educational level of the fathers and vocabulary skills (Mill Hill vocabulary test) of each parent, and their RF scores (see Table 3).

	Level of education	IQ	social class	socio-economic group	age	years of cohabitation
mothers	.19	.33*	-.11	-.04	.12	-.10
fathers	.35*	.27*	-.24	-.01	.15	-.02

Table 3. Correlation coefficients between RF ratings and demographic characteristics of 100 mothers and 100 fathers

We computed the correlation coefficients between RF ratings and other ratings derived from the Adult Attachment Interview. The correlations are shown in Table 4 below.

AAI scales	Mothers' AAI (n=100)	Fathers' AAI (n=100)
Loving relationship with mother	.426	.425
Loving relationship with father	.439	.279
Role reversal in parental relationships	-.298	.130
Poor quality of recall	-.633	-.246
Current anger with parents	-.123	.20
Idealisation of relationships	-.424	-.214
Derogation of relationships	-.349	-.253
Coherence	.734	.638

Table 4. Correlations between AAI scales and RF ratings

In analysing the scores given to Adult Attachment Interviews on the original scales, and some additional scales including RF ratings, we computed a correlation matrix based on the 200 interviews. 6 factors emerged from the 29 scales. These were: 1. Negativity in relationship to mother; 2. Negativity in relationship to father; 3. Coherence in Internal Working Models of attachment; 4. Overprotection; 5. Pressure to achieve; 6. Guilt. Each had a coefficient alpha above 0.8, and the Coherence factor had an alpha of .9 for women's AAIs and .92 for men's AAIs. Factor loadings are shown in Table 5.

	Factor loadings	Correlations of ratings with factor total
Reflective function	.91	.93
Coherence of mind	.77	.83
Subjective experience of poor recall	-.74	-.76
Objectively poor recall	-.73	-.71
Idealisation of mother	-.69	-.45
Coherence of transcript	.67	.74
Passivity of thought	-.51	-.63
Total eigen value	2.59	-

Table 5. Loadings of scales on Coherence factor, and correlations of each scale with factor total (n=200)

In order to investigate the unique contribution of each of these scales to the Coherence factor score, a stepwise multiple regression was carried out, with the factor score as the dependent variable and the scale scores as independent variables. RF was entered first, as the highest loading variable, and accounted for 86% of the variance. The coherence of transcript score accounted for a further 6% of the variance, and objective poor recall an additional 5%. The association between AAI classification and this Coherence factor, dominated by the RF scale, was strong. Autonomous (Secure) classification was associated with higher factor scores on Coherence ($F=62.99$, $df\ 2,197$, $p<0.0001$). There was little difference between the two insecure classifications. The point biserial correlation between security-insecurity on the Adult Attachment Interview and the Coherence factor score was 0.75. Overall, the psychometric analysis of the AAI, as reported in Steele, revealed that ratings on RF were consistently the strongest contributors to judges' assessment of attachment security, and accounted for more than half of the variance in the secure/insecure distinction.

There was a strong relationship between scores on the RF scale and the Strange Situation behaviour of infants, whose mothers and fathers had been assessed using the AAI before the birth of the child (see Table 6). The point biserial correlation between secure classification in

infancy and parental RF was highly significant ($r=.51$ for mothers and $.36$ for fathers, $p<.001$ in each case). The correlations observed with ratings of the infants' behaviour in the Strange Situation were also significant. These correlations are shown in Table 7.

		Infant Classification		
		Avoidant	Resistant	Secure
No reflective function	M	4 (1.6)	0 (.6)	1 (2.8)
	F	3 (1.4)	0 (0)	2 (3.6)
Poor reflective function	M	18 (11.8)	4 (4.2)	15 (21)
	F	18 (12.7)	0 (0)	28 (33.3)
Average reflective function	M	9 (14.7)	6 (5.2)	31 (26.1)
	F	3 (9.7)	0 (0)	32 (25.3)
Good reflective function	M	0 (2.9)	1 (1)	8 (5.1)
	F	0 (.3)	0 (0)	1 (.7)

Table 6. Frequency distribution (expected frequencies in parentheses) of infant attachment classification according to parents' RF scores (M = mothers; F = fathers)

Ratings of infant behaviour				
	Proximity seeking	Contact maintenance	Resistance	Avoidance
Mother (n=97)	.255**	.300**	.020	-.375***
Father (n=87)	.140	.167	-.124	-.235**

Table 7. Correlation between parents' RF ratings and infant behaviour in the Strange Situation

In a subsequent study on the same sample, we found that RF was particularly predictive of secure attachments with mothers, in cases where mothers independently reported significant deprivation in childhood (Fonagy et al., 1994). For the purposes of this analysis, we divided the sample of mothers into those with high and low deprivation in childhood. High deprivation was defined as showing three or more of the following indicators: prolonged separation from parents before 11 years, prolonged paternal unemployment, life-threatening illness of either parent, psychiatric illness in either parent, boarding school, overcrowding, single parent family, serious illness in the child, low socio-economic status. Mothers were divided into those showing low or high RF (score over 5) ($n=48$ and 49 , respectively). Of the non-deprived group, 79% of those with high RF ($n=39$) had secure infants compared to 42% of those with low RF ($n=31$). In the deprived group, 100% of those with high RF ($n=10$) had secure infants, whereas 6% of those with low RF ($n=17$) did so. These results show that RF is particularly important in facilitating attachment security between infant and mother where the mother has suffered social deprivation in childhood.

As part of establishing the discriminant validity of the Reflective Function Scale, the scale was related to a number of psychometric instruments. Epstein's Mother-Father-Peer Scale (Epstein, 1983) measures independence-encouraging versus over-protective and accepting versus rejecting mothers, fathers and peers, none of these scales related to RF scores (see Table 8). The Eysenck Personality Questionnaire (Eysenck & Eysenck, 1975), concurrently administered to the subjects, showed no relationships to extraversion, neuroticism or psychoticism. The Langner 22 (Langner, 1962), a screening measure for psychiatric caseness, also showed no correlation with RF ratings. The RF scale was also correlated with the Sources of Self Esteem Inventory (O'Brien, 1981). The RF Scale had no significant association with any of the 11 scales of the SOSE.

	Scale	Correlation with RF	probability
Epstein's Mother-Father-Peer Scale	independence vs overprotection, mother	.05	n.s.
	independence vs overprotection, father	-.11	n.s.
	acceptance vs rejection, mother	.02	n.s.
	acceptance vs rejection, father	.19	n.s.
	acceptance vs rejection, peer	.11	n.s.
Eysenck Personality Questionnaire	psychoticism	-.11	n.s.
	extraversion	.21	n.s.
	neuroticism	-.03	n.s.
	lie scale	-.21	n.s.
Langner 22		-.11	n.s.
Sources of Self Esteem Inventory	global self-esteem	.06	n.s.
	competence	.10	n.s.
	loveability	.13	n.s.
	likeability	.05	n.s.
	self-control	-.03	n.s.
	personal power	.12	n.s.
	moral self-approval	.02	n.s.
	body appearance	-.01	n.s.
	body functioning	-.02	n.s.
	defensive self-	-.11	n.s.

enhancement integration and identity	.08	n.s.
--	-----	------

Table 8. Correlations between RF ratings and questionnaire measures of past relationships and current self-esteem (n=200)

In order to assess whether mood state immediately prior to or immediately following administration of the AAI might have influenced RF rating, a previously validated single item affect grid was employed (Russell, Weiss, & Mendelsohn, 1989). There was no relationship between mood, either before or after the interview, and RF rating.

3.2 The Cassel Hospital Study

82 out of 85 consecutively admitted non-psychotic inpatients at the Cassel Hospital participated. The hospital is a UK national centre for the inpatient treatment of severe personality disorder. Patients at the hospital were rarely on medication at the time of the study. The group were matched on age, gender, socio-economic status and verbal IQ with 85 normal control participants recruited from an outpatient medical department. In both groups, the average age was 29, 15% of the sample was male, 50% were from social classes I and II; mean Verbal IQ was 114.6 for the patient group and 112.1 for the control group. Diagnoses were made on the basis of the SCID-I and SCID-II (DSM-III-R).

The results of this study have been reported elsewhere (Fonagy et al., 1996). Here only data concerning RF ratings will be reported. The Adult Attachment Interview was administered to all patients and controls and coded for RF by two raters. The agreement between the raters was .91. Mean RF ratings for the psychiatric and control groups were 3.7 (sd=1.8) and 5.2 (sd=1.5) respectively ($p < .0001$). The mean RF ratings for the common Axis I and Axis II diagnoses are shown in Table 9 for the psychiatric group only. The Axis I diagnoses did not distinguish high and low scorers on the scale with the exception of eating disordered patients, many of whom also carried an Axis II diagnosis (particularly BPD). Patients without Axis II diagnosis were rated higher on RF than those with ($p < .05$). This was principally due to the low RF scores of patients with a diagnosis of BPD ($p < .001$).

		Diagnoses		
Axis I	Depression (n=72)	Anxiety (n=44)	Substance Abuse (n=37)	Eating disorder (n=14)
	3.8 (1.7)	3.5(1.8)	3.4(1.8)	2.8 (1.7)*
Axis II	No Axis II (n=23)	BPD(DSM)	Antisocial or Paranoid (n=22)	Other (n=38)
	4.3 (1.7)*	2.7 (1.6)***	3.9 (1.8)	3.3 (1.7)

Table 9. Means (and standard deviations) for RF ratings of 82 psychiatric patients grouped according to diagnoses (Asterisks indicate significant differences between groups diagnosed with a disorder compared with all those in the psychiatric group not diagnosed with that disorder. Means are adjusted controlling for age, Global Assessment of Functioning scores on admission).

We examined the association between reported physical and sexual abuse, RF scores and a diagnosis of BPD in a 2 X 2 X 2 hierarchical log-linear analysis using the score of 3 as a median split for RF. Thirty-two of the 53 (60%) cases who reported abuse were independently diagnosed with BPD, compared with 4 of 29 (14%) who did not report abuse. The likelihood of reported abuse being associated with BPD was greater in the group of patients with low RF than those with RF ratings above the median. Only 4 of 24 (17%) patients reporting abuse in the high RF group were diagnosed with BPD, whereas 28 of 29 patients (97%) reporting abuse in the low RF group reporting abuse were so diagnosed. In the group not reporting abuse the prevalence of BPD was the same in low and high RF groups (2 of 17 for high RF vs. 2 of 12 in low RF). Thus RF is predictive of BPD only in the presence of abuse. In line with this argument, the three-way interaction component of the log-linear analysis was significant (chi squared = 8.67, N=82, $p < .004$).

3.3 The Prison Health Care Centre Study

This study involved 22 prisoners on remand referred to the Health Care Centre of a local remand prison for psychiatric diagnosis. All subjects were administered the AAI along with a range of diagnostic and psychometric instruments. Prisoners were only included in the study if they also had a diagnosis of a least one Axis I or one Axis II disorder. In the event, SCID I and SCID II (DSM-III-R) interviews identified three or more Axis I diagnoses in 80% of the sample and at least one Axis II diagnosis in 91%. A diagnosis of BPD was made on 11 of the 22 cases. The offences ranged in severity from theft, burglary and damage to property, to gross indecency, grievous bodily harm, malicious wounding, multiple armed robbery, kidnapping, rape and murder. The offences were later subdivided into less violent (against property) and more violent (against person). The mean age of the group was 28 (21-40), mean IQ 110 (80-126), modal social class was IV (III-V).as classified on occupations before imprisonment, all subjects were white except for three (2 West Indian and one African). There were two comparison groups. A personality disorder group was recruited from a hospital which is a national centre for the treatment of personality disorder. An attempt was made to match the groups on diagnoses which was quite successful for most diagnostic categories except anxiety, which was more common in the hospital group. The severity of the disorder (GAF score) was also significantly lower in the Hospital group (48 vs 29). The two groups were well matched on Axis II diagnoses, age, IQ and socio-economic status. The third group, recruited from a medical outpatient department was screened for psychiatric disorder and matched for gender, age, IQ and socio-economic status.

The mean RF ratings of the prison group was 2.5 (sd=1.8). This compared unfavourably with the scores obtained by the more impaired inpatient group of 3.7 (1.5) and the normal control

group of 5.8 (2.3). The difference between the inpatient and prison groups was significant ($p < .01$). The proportion of the prison sample scoring less than 3 was 64% compared to the inpatient group of 23% ($p < .01$). There was an insignificant trend for patients with BPD diagnosis in the prison group to have RF scores below the cutpoint of 3. 83% of BPD prisoners fell below the cutpoint but only 54% on prisoners without a BPD diagnosis did so ($p < .10$). A greater proportion of violent criminals had RF scores below the cutpoint than non-violent criminals ($p < .05$). These results are part of a pilot investigation (Levinson & Fonagy, submitted) and the sample size is too small to draw definitive conclusions. The RF scale nevertheless appears promising in distinguishing criminal groups with mental disorder from groups with similar disorders but without criminal tendencies.

3.4 The use of reflective function ratings in predicting the child's capacity to mentalize

This study examined the relationship between mother's RF scores obtained on the bases of interviews before the birth of the child and the child's performance in a task which involved belief-desire reasoning (Harris, Johnson, Hutton, Andrews, & Cooke, 1989). The task requires the child to identify accurately the emotion state of a puppet based on the false-belief of the figure. Most children pass this task by 6-7 years of age. In the current sample (described in section 3.1) children performed the task at 5.5 years and about two thirds passed the test. RF scores of the mothers were used to predict the child's performance.

Mother's RF, assessed on the basis of the prenatal interviews, was found to be highly predictive of the child's success in the Belief-Desire Reasoning Task ($r(90) = .32$, $p < .001$). This correlation controls for both the child's and the mother's verbal ability. In a path analysis, which included mother's attachment security, father's attachment security, mother's RF (metacognitive ability), infant-mother and infant-father attachment security, and child's verbal fluency as predictors of the child's performance on a cognitive emotion task, mother's RF was found to predict the child's performance, both via its influence on the child's attachment to the mother, and directly (see Figure 1). This suggests that the impact of mother's RF on the child's development of mentalisation is not simply through the enhancement of secure infant-mother attachment, but it may also influence other aspects of the mother-child interaction, including perhaps the nature of conversations, the amount of pretend play, the nature of disciplinary interactions, all of which have been shown to enhance theory of mind competence in children (Dunn, 1996; Fonagy, 1997b).

4 ILLUSTRATIONS OF MODERATE TO HIGH RF

The examples below indicate instances where mentalization is explicit in attachment-related narratives. There are three considerations which should be taken into account in identifying these instances. First, it is most unlikely that any specific instance will be a "pure" example of any one category, and it is not necessary to identify all categories shown in a single sentence (however, at the level of rating passages, described later, the number of categories included in the passages is important to note). Second, the example should originate from an attachment context, which includes self-other interactions and the impact of self-states in one's own or

other behaviour. Third, statements must be specific in two ways: a) they must be relevant to the situation which is being described, rather than be a generalisation about larger categories of situations. Thus, general comments about school, home, family or work are less likely to be indicative of Reflective Function than comments about specific behaviours of individuals in those contexts; b) they must be specific to mental states, rather than personality dispositions, such as friendliness, negativity, anxiety. There will be examples of these considerations later. In general, the more detailed and elaborate the mental state description the more likely it is to be an example of reflective function.

4.1. Awareness of the nature of mental states

Raters are asked to note the interviewees' awareness of the characteristics of mental states in themselves and in others, by making references to one of the characteristics of mental states (such as the opaqueness of mental states, their frequently self-serving bias, or the fact that thinking or wishing something does not make it true). Below is a list with illustrated example of these and other characteristics of mental states. The list is not exhaustive or mutually exclusive. It would be impossible to come up with an exhaustive list of examples because one can never completely know one's own or another's mind; and it would be impossible to provide a list of discrete mental state characteristics because speakers' references to mental states almost invariably refer to more than one, and often many, of the characteristics listed below. In the examples offered, when a given example may be ambiguous as to which characteristic of mental state understanding it depicts, the portion of the speech sample thought to be most relevant to the characteristic being discussed is underlined, and we recommend the same procedure for rating.

4.1.1 The opaqueness of mental states

The speaker acknowledges the difficulty one has in being sure of what the other's intention or mental state is or was, while being prepared to guess. Thus, the statement "*I thought my mother felt resentful of us, but I'm not really sure if she felt that way herself*" (5) would be regarded as reflective, whereas the statement "*One can never know what anyone else thinks*" (1) would not, without a suggestion as to what may have been thought. Such an explicit clue is rare, and indications of opaqueness usually emerge through qualifiers such as "perhaps" or "might", prefacing an explanation in terms of mental states. Even more convincingly, alternative mental states may be offered, with the implication that it is hard to know which one lay behind the behaviour.

4.1.2 Mental states as susceptible to disguise

Related to the issue of opaqueness is the possibility of deliberate disguising of internal states. Recognition of this possibility may be implicit or explicitly stated. A common example would be instances of awareness that the individual may experience different emotions to the ones they display, and may refer to the other or to the self (e.g. "*I am so angry at her ... but I would never show that to her.*" (4)) "*My mother always kept everything controlled and calm, but I think underneath she often felt very angry*"

4.1.3 Recognition of the limitations on insight

Another sign of awareness of the nature of mental states is the explicit qualifying of insight concerning oneself or others, i.e. awareness of one's limitations in being able to understand self and others: *"I had a lot of respect for my mother, although sometimes she infuriated me because she was a very, er, anxious person, and, um, she would get very uptight about things and sometimes get a little bit hysterical, I think she was very insecure in her relationship with my father, but I don't know if that was true"* (6) *"I had a real go at her, but I think I was just frightened that she didn't care about me."* (6)

4.1.4 Mental states tied to expressions of appropriate normative judgements

Mentalising capacity is indicated by awareness of an expectable psychological response. When an individual refers to what would be a commonly expected reaction in a specific situation, this is rated as evidence of reflective-functioning, e.g. *"I started staying out at night and he became quite authoritarian like fathers do, but I think all fathers get anxious about their daughters when they sort of - start rebelling."* (4) There is often a sense of forgiveness implied by such statements, but this aspect is not rated, e.g., *"It was understandable that my mother would feel jealous of me because of my close relationship with my father."* (4)

4.1.5 Awareness of the defensive nature of certain mental states

Explicit recognition of people's tendency to modify their mental states in order to reduce negative affect is rated as a manifestation of the RF. For example: *"You tend to blank out things that make you unhappy sometimes"*. (4) This can be applied to self and others.

4.2 The explicit effort to tease out mental states underlying behaviour

A further illustration of moderate to high RF appears when speakers identify possible mental states which may account for behaviour, and offer accurate or plausible links between mental states and behaviours of the self and others. This is evident in a range of statements illustrated below.

4.2.1 Accurate attributions of mental states to others

Raters should note all plausible causal accounts of behaviour in terms of mental states. For example: *"Mum would say 'Wait till your Dad comes home!' Then he would wallop us. But, you know, I think he felt obliged to do that because he knew that she had to be with us all day, he felt he had to back her up."* (6-7). The rater should be careful to give credit only for specific rather than generalised, and common-sensical rather than theoretical attributions. The aim here is to distinguish genuine instances of mentalization derived from experience from understanding derived from external sources such as self-help books.

4.2.2 Envisioning the possibility that feelings concerning a situation may be unrelated to

observable aspects of it

This can be seen when a person recognises that affect is inconsistent with the external situation. The following examples would be credited with reflective-functioning *“I’m looking after my daughter for three weeks and it’s .. it’s .. er ... I’m finding it very hard. But I’m only finding it hard because I’m a bit unhappy with certain other aspects of my life at the moment, apropos my career ...”* (5) *“We moved into a new house, and Mum and Dad were disappointed because I hated my new room. It was bigger and lighter but I loved my little room in Street X.”* (3-4)

4.2.3 Recognition of diverse perspectives

The speaker explicitly acknowledges that different people may perceive a given behaviour or situation differently. For example: *“My mother had the habit of lifting her hand and slapping us, um, if we were naughty, or when she thought that we were naughty.”* (5) My father thought it was fine for Mr X (the teacher) to behave like that, that it would teach us self-discipline and that kind of stuff, my mother thought that it was appalling to treat children that way. I think he just didn’t know what was going on. How all the children felt” (5)

4.2.4 Taking into account one's own mental state in interpreting others' behaviour

Speakers who make explicit reference to how an interpretation of an event might have been distorted by what they were feeling or thinking at the time should be credited with this understanding, e.g., *“I saw him as behaving in an uncaring way but actually I was very angry with him at the time because of the way he treated my mother ... So perhaps, that’s why.”* (6)

4.2.5 Evaluating mental states from the point of view of their impact on behaviour of self and/or other

Speakers recognise the role their own mental states might have had on the behaviours of self and others. Implicit, but not necessarily explicit, is that the mental state would have had behavioural manifestations. For example: *“Because I think perhaps I was insecure, a very demanding child that needed a lot of attention. I obviously upset her by being so demanding and crying whenever she went out .”* (6)

4.2.6 Taking into account how others perceive one

Speakers show explicit awareness of how others’ perception of them is related to their own or others’ actions and reactions. For example: *“People always saw me as somewhat haughty and full of pretensions, but really I just felt very insecure. I realise now that it put people’s backs up.”* (6-7)

4.2.7 A freshness of recall and thinking about mental states

Speakers give the impression of thinking spontaneously and vividly about people’s thoughts and feelings. This is the opposite of merely learned or clichéd expressions (See 5.2 below). There is

a quality of something currently thought, and real to the subject, which makes it feel alive to the rater. This will often be conveyed, for instance, when a subject changes their perspective on an event or relationship during the interview itself, and is often marked by dysfluency, as the subject struggles to formulate a new understanding. For example: *“I still feel the same way about him dying, it was cruel and unfair for someone so young and with so much more to give to die, ... so I don’t think my feelings have changed, ... But now -- as I think about it -- I realize, that in fact I hardly ever think about him or his death these days, and I actually, well, probably resent it a bit when other people keep talking about him.”* (6)

4.3 Recognising developmental aspects of mental states

A feature of speech that provides convincing evidence of moderate-to-high reflective-functioning is an explicit reference to developmental aspects of mental states. These include acknowledging the influence of one generation upon the next, showing an understanding of how mental states of others’ change, showing an appreciation of family dynamics, and distinguishing between the thinking of a young child and older person. These and other developmental aspects of mental states are illustrated below.

4.3.1 Taking an intergenerational perspective, making links across generations

Basic to attachment theory and other psychoanalytic approaches is the assumption that parenting behaviour is fundamentally influenced by parents’ thoughts and feelings regarding their childhood relationship experiences. Statements showing awareness of this intergenerational exchange of ideas, feelings and behaviour is considered reflective as long as the references made are explicit and specific. Thus, while the following statement would count as reflective *“My mother expected great things from her children because, as I’ve been saying, so much had been demanded of her by my grandparents and, at least in my mother’s mind I believe, she never lived up to those unrealistically high expectations, and so hoped that we would”* (5), the following example on the same theme would not count: *“What influenced my parents to be like that? I think it was their parental upbringing. I think my grandmother who brought up my mother with the aid of her mother, I mean it goes back a long way these mothers and daughters in our family ... I’m sure my mother is the person she is because of the way she was brought up. I’m sure my grandmother was the greatest influence on her ...”* (2) If the subject had mentioned (as in the first example) that the grandmother was very ambitious for her daughter and this explained to the subject her own mother’s inclination to push her towards achievement, then the example would count.

4.3.2 Taking a developmental perspective

Some subjects show awareness of developmental changes in certain mental states. This is regarded as reflective because it assumes that the subject is making assessments of either their own or others’ changing perspective with age. For example: *“When we were little my father always seemed to have time for us and we would have so much fun together, but then as we got*

older he withdrew and had difficulty I think getting on with teenagers.” (4)

4.3.3 Revising thoughts and feelings about childhood in light of understanding gained since childhood

Views of the world, feelings and beliefs concerning it, change radically between childhood and adulthood. Children’s understanding of the social world is particularly limited. Some individuals show awareness of the implications of such changes for changes in their behaviour or attitudes. These are considered reflective. *“It’s only now in adulthood that I realise she was ill. When I was a child I thought of her as either just not liking me much or as very withdrawn and shy. Now I can feel for her much more and she does not make me sad or angry anymore.” (6)*

4.3.4 Envisioning changes of mental states between past and present, and present and future

A key indicator of reflective-functioning is manifested when speakers consider developmental changes in their own mental states. The change from childlike to adult ways of thinking, we have dealt with above, is one major indication of this kind of conceptual framework. Changes however may be envisioned over a much shorter time-span, particularly when considering the future. For example, consider this statement from an expectant mother contemplating the wishes she has for the baby son she is carrying: *“Now I just want him to be healthy and happy with what he has, but I know once he’s born I’ll want him to be Prime Minister.” (5)*

4.3.5 Envisioning transactional processes between parent and child

A full understanding of parent-child behaviour is not possible without taking into consideration the impact of the child on the parent. The transactional perspective is considered by us as necessarily involving reflective function. For example: *“I think not only did my mother hold on to me very tight, but I wrenched away very hard, and those two actions made it much more extreme than it might otherwise have been” (5) or “I think my shyness was upsetting and painful for her and as a result I irritated her. I think that’s why she avoided me.” (6)*

4.3.6 Understanding factors which developmentally determine affect regulation

A parent’s awareness of a child’s emotional needs requires reflection on the mental world of the child. For example, the small child is unable to regulate his own emotional state and is dependent on the caregiver to perform this homeostatic function. Soothing is a good example of the caregiver’s attempt to regulate, in this case to reduce, the infant’s arousal. Awareness of the mental state of the baby, and his helplessness vis-à-vis his own emotional state, is considered an instance of reflective-functioning even if this is apparently about a child unrelated to the subject. For example: *“I saw that he (boy of 4) just needed a bit of reassurance, he was being naughty to test his parents’ love for him.” (5)*

4.3.7 Awareness of family dynamics

The speaker shows an awareness of the interdependence of mental states within family systems. Seeing the family as an interdependent system, where the mental states of the individual members interact and create attitudes and feelings which each individual member is affected by, indicates a high level of reflection. These examples are rare, but when present, are highly compelling. For example: *“Was anybody rejecting? Not really. It is funny ... I remember both my parents as very loving and I think they were, and my sister and I got on extremely well. Yet if you asked me to describe the family, I would say it was cold and somewhat rejecting. Somehow, perhaps because we tried not to show preferences, we all managed to make each other feel as if we liked somebody else more. So we all managed to feel a bit rejected and unloved.”* (7)

4.4 Mental states in relation to the interviewer

A subject's recognition of mental states might be shown by their interaction with the interviewer, which we take as an indication of the subject's willingness to entertain mental states in the context of other relationships. This tendency may again be revealed in a number of ways:

4.4.1 Acknowledging the separateness of minds

The separate history, experience and mental stance of the interviewer is at times explicitly acknowledged by the speaker and this provides a compelling example of high reflective-functioning. Speakers demonstrating this understanding make explicit their appreciation that the interviewer may not necessarily share their mental state. For example, one speaker said on becoming upset in an interview some time after having alluded to a traumatic road traffic accident happening some time in the past, *“It must seem strange to you that I'm still upset, but it is almost exactly this time of the year when the accident happened.”* (3) One person, claiming to have felt closer to her father, whom she rarely saw, said: *“I suppose that sounds odd, but actually it is easier to feel close to someone you don't see a lot, in some ways. Closeness brings all sorts of fears, of losing that person or being hated. If you don't see someone they can't hurt you!”* (6)

4.4.2 Not assuming knowledge

Most commonly it is speakers' lack of consideration of the mental state of the person they are talking to which is easiest to observe and rate. Some subjects clearly assume knowledge which they should know the interviewer has no access to. In some instances a protagonist is referred to by a newly mentioned nickname or relevant history is assumed or the interviewee gets so absorbed in their own thoughts they appear to forget that they are recounting a story to someone. All these instances should be noted and taken into consideration in giving a final rating. Explicit efforts by the subject to help the interviewer keep track of the material by stepping outside the narrative and spontaneously clarifying confusing aspects should be credited. This is clear when a subject acknowledges ambiguities or anomalies in a narrative, and provides additional information which either clarifies or explains. For example, a woman who was brought up in a very different culture, with several “primary” caregivers, in an extended family said that she very rarely approached her father for comfort. She acknowledged that the

interviewer would find this surprising, and explained that *“In Country X, the children always go to the grandmother or aunt, they are not allowed to bother the father or uncle.”* (4) Less persuasive examples are when a speaker pauses to ask whether the interviewer is following the narrative. *“Is this still relevant to the question?”* or *“Have I got off the question?”* (up to 3) This should only be the case, however, when, in the rater’s view, the subject is accurately and selectively responding during particularly complex sections of the narrative. Constant clarifications for the sake of the interviewer, may reflect a defensive style of narration rather than actual concern for the interviewer’s level of knowledge.

4.4.3 Emotional attunement

Speakers should also be credited for making accurate references to the likely impact upon the interviewer of the material they have provided. Some individuals, even when upset themselves by their recollections, assume that the interviewer is unaffected by what they hear. Others are far more attuned to the interviewer’s likely emotional reaction, and indicate their awareness in their narrative. The rater must take care not to mistake common courtesy or a wish to please for such attunement. Phrases such as *“I am sorry this must be so boring for you”* do not count, even if they are accurate. It is important that the subject should manifest an internal model or a hypothesis as to why the interviewer may be upset, bored, irritated, frustrated etc. For example one interviewee, on noting that the female interviewer was somewhat upset by the account of a cot death of a sibling, momentarily interrupted the narrative to acknowledge that: *“To lose a child like that, without reason is horrible and even hearing about it can be upsetting, it must be quite difficult to listen to it”*. (4) More commonly, accounts of severe maltreatment or abuse impact deeply upon the interviewer. Reflective subjects may explicitly acknowledge the likely effect their story is having on their listener: *“I realize that it can’t be nice to have to hear about this. I wish I had a nicer story to tell”*. (3)

5 GENERAL CONSIDERATIONS

There are a number of considerations which should guide the rater’s judgements as to whether or not a given statement indicates RF. These are outlined below.

5.1 Only explicitly reflective statements qualify for high ratings

People commonly refer to mental states as part of their discourse. Commonly, such statements simply imply language competence, rather than a true consideration of mental states. For example, children use mental state terms long before they acquire an explicit “theory of mind” (Bretherton, Fritz, Zahn-Waxler, & Ridgeway, 1986). We habitually say *“I think ...”*, very rarely implying that we have actually thought about the issue. Thus a simple count of mental state terms is inadequate for the operationalisation of mentalising. Therefore, for a statement to be reflective, it has to imply an activity of reflection, rather than simply describing a mental state. Thus, the statement *“I was angry”* would not be considered reflective, whereas the statement *“At the beginning I felt really angry and then it changed to sadness”* (5) does suggest that the feelings have been considered. Raters should attempt to identify instances where there is thinking about a belief or a feeling, as opposed to a simple report. However, a simple

description can actually contain reflection, provided that it is elaborated in such a way as to make a link between the state of mind and their own or another person's behaviour.

5.2 Learned, rote, or clichéd statements do not qualify for high ratings

Subjects frequently make references to mental states and appear to reason with mental states (i.e. place them in the context of causal explanations), yet the statements sound parroted or clichéd. These should not be rated because they could have been learned by rote as entire phrases and need not reflect the speaker's specific understanding of meanings in mental state terms. There are exceptions to this general rule, specifically, where learned or familiar statements are sufficiently supported by specific descriptions of the context to make the rater confident that the subject is using a commonly used expression or phrase to describe a personal understanding. For example: *"I mean I think my mother was preoccupied with her own infantile needs and she couldn't really get too far beyond that"*- which on its own would not be considered reflective, but later in the interview this was supported with accurate descriptions of the mother's childlike neediness and the limitations this placed upon her ability to parent the child. Statements such as, *"I was the Oedipal son"* or *"My father and I rivalled for my mother's affections"* would not count as reflective; without further elaboration and convincing evidence they should be considered clichéd and non-reflective statements.

5.3 Reference to personality or a relationship, in the absence of specific reference to mental states, does not qualify for a high rating

Narratives frequently contain justifications of behaviour in terms of somebody's personality, psychiatric condition, or the nature of a relationship. While in some circumstances this may indicate reflection, we think that there is an equal number of cases where these derive from phrases drawn from folk psychology, in other words they are generalisations of culture. For example, *"My mother and I are very similar, we understood one another all too well, that is why we never got on."* (3-4) would be seen as only mildly reflective, whereas the following which roots the generalised comments about parent personality to context would be considered reflective: The speaker was asked why she thought her parents behaved as they did during her childhood, *"I think my parents both withdrew from one another and from us because both of them expected a different relationship. Mother wanted a husband who was there and interested, and my father thought marriage was going to the pub while his wife cleaned up and cooked. So in pursuit of their separate visions of family life they rowed a fair bit, and mostly avoided spending much time with us."* (6)

5.4 Avoiding thinking for the subject

It is very tempting to attribute meaning in terms of mental states, filling in the detail or examples which the subject has not in fact provided, even if they have spoken about examples earlier. This giving of the benefit of the doubt must be carefully avoided, for example when a subject offers a plausible generalisation. A subject might explain that his relationship with his parents got worse after he was sent away to school, and that he never felt the same after that. It would be easy to

supply the feelings of sadness, rejection, resentment etc, but these have not actually been stated, and cannot be inferred.

5.5 Diagnoses should not be accepted as a shorthand for mental states

The use of diagnostic terminology, or reference to mental illness, should be considered very carefully, and on the whole rated low, if this is the sole explanation offered for the caregiver's behaviour, and the specific mental states of caregiver and other persons affected are not specified. Very commonly, subjects refer to depression in the mother as an explanation for neglect or rejection. This should not be scored as reflective unless the subject also supplies the perceptions and beliefs on the part of the child and/or parent associated with the behaviour attributed to the psychiatric condition.

6 ILLUSTRATIONS OF NEGATIVE OR LIMITED RF

6.1 Possible implications of sub-types of impoverished reflective-functioning

We have not as yet arrived at a typology of good reflective-functioning. Although there may be several styles of mentalizing and these may be important in predicting other behaviours we have no reliable scheme for distinguishing these. In our experience, the different aspects of mentalizing noted in this manual tend to occur together to a greater or lesser extent. Perhaps independent scoring of each of the above aspects of reflective-functioning may eventually lead to the identification of patterns or combinations of different types of capacity, but at present we do not advise using separate scores.

We have however noted that when reflective-functioning is absent or ineffectual, quite distinct, readily classifiable, patterns of responses tend to emerge. At this stage we have only limited evidence that these patterns are of psychological significance. They may be conceived of in two ways: 1) they may be thought of as different specific interferences with reflective-functioning which may take on importance in the same way that choice of defences seems to be relevant in other contexts; 2) the type of limitation on RF may relate to underlying relatively stable personality dispositions or disorders, or more transient psychiatric conditions. For example, a person with a paranoid illness may show a rejection of reflectiveness, which is part of the psychiatric state.

If the former explanation is the case, the typology outlined below indicates different ways in which individuals avoid mentalisation or different components of the process of mentalisation which has been compromised by developmental processes, psychic conflict or limitations of the psychosocial environment. If the latter explanation is the case, the typology offered below may relate more to other aspects of personality functioning including patterns of attachment, and may not in fact illuminate our understanding of reflective processes beyond assisting raters in arriving at an attachment classification. Of course there may be a third, rather trivial reason for the existence of such distinct patterns. Speakers, asked by the interviewer to reflect on their experiences but unable to do so, must find some means through which they can meet the demand characteristics of the situation. It is possible that the different groups of interviews

reflect different coping strategies which the person arrived at for dealing with the specific situation in that specific setting and the categorisation has very limited psychological significance.

The categories of absent, low or questionable reflective-functioning listed below are offered tentatively, in the hope that further research may help clarify the meaning of these individual differences in responding to the demand to reflect upon and evaluate the meaning of one's attachment history. We suggest the typology should be used alongside all low ratings (up to '3') on individual particularly striking passages rated, as well as in applying one overall score to the transcript. This process may be helpful in assisting raters to arrive reliably at an overall category for the entire transcript, as the different examples of negative or absent RF may be used to indicate the likely category of low rating to be assigned to the interview as a whole. Here the subtypes are listed from least reflective to more reflective subtypes.

6.2 Rejection of RF

Some individuals become overtly defensive about questions having to do with mental states underlying their own and/or others' actions. They may perceive the interview situation as an attack, and become quite hostile in response to mild interrogation. Interviewer: *"Why do you think your parents behaved as they did?"* Subject: *"How do you expect me to know? You tell me, you are the psychologist!"* (-1)

A second aspect of these transcripts is the frequent or total evasion of questions related to mental states. Such evasion is specific to the interview context where thoughts and feelings with regard to the primary caregivers are explored and may not characterise the rest of the interview. *One subject, when asked why he thought his parents had behaved as they did, told the interviewer about how having been exposed to football (soccer) early in life had developed a life-long devotion to the game.* (-1)

In general, such interviews tend to be relatively short, the relationship with the interviewer is business-like or overtly hostile, the task of having to think about childhood is resented and unequivocally not enjoyed, there is no sense of co-operation between the interviewer and subject, misunderstandings may abound, and mention of internal states (feelings, desires and wishes in particular) may be completely absent. The mental state language of the subject, if any, may be idiosyncratic, and the rater may have the feeling that the subject just does not understand what the interviewer means when he/she refers to internal states.

6.3 Unintegrated, bizarre or inappropriate RF

A rare but interesting category of low reflective-self functioning refers to a fairly heterogeneous group of strategies towards reflection in the interview. What transcripts classified into this category have in common is the difficulty experienced by the rater in understanding or following the subject's attributions of mental states to themselves or others. The difficulty may arise from a number of disparate sources, but these may be broadly categorised under two headings: a) failure of adequate elaboration, and b) interference from inappropriate cognitions.

We shall describe each subcategory in turn.

Some subjects disregard the inherent interdependence of various sorts of states of mind. Most commonly subjects appear not to recognise that affective states are likely to be underpinned by either beliefs or behaviour. For example, a subject may mention intense negative affect towards a caregiver without giving reasons for it. (For example, *"I ... I just disliked her"*.) The rater (and sometimes the interviewer) may feel a justifiable need to ask 'why'. In other instances the reason may be given but the transcript does not allow the rater to make an appropriate inference as to affect or behaviour. The different statements seem not to hang together. For example, the subject may express a belief that the caregiver made extraordinary effort to be present at important times in the subject's life, yet may ignore this evidence in coming to a generalisation about the caregiver's emotional stance. For example, in one interview the subject described how his mother took time off work every time he was ill, turned down opportunities for promotion in order to be available after school, took all holidays to ensure her availability at times of transition or other crises; yet the subject said: *"I don't know how she felt about me. That's one of the big mysteries of my life. She never said what she felt towards any of her children."* (0)

This tendency, which appears to be a lack of integration, creates a sense of lack of meaning surrounding mental state attributions. Mental states may be stated explicitly but the implications of ideas, beliefs and feelings are not spelled out, or are not reasonable. Intense feelings may be attributed to the other and presented as if they had no meaning for the subject; nothing is said about how it makes them feel. For example: *"She was a curious person, my mother, unpredictable, even odd at times. One day she cut down all the flowers in the garden ... The other thing was that she was very shy ..."* (-1) Also commonly the parents' relationship may be described but its implications for the child are the reverse of one she would expect. For example, the parents' relationship may be described as particularly distant and cold and amply illustrated as containing much resentment and passive aggression, yet the interviewee describes the family as being very happy.

In many interviews elaboration may only be achieved when the interview is considered as a whole. Thus passages should only be identified as reflecting a lack of integration if the rest of the interview offers no elaboration or substantiation.

A related but distinct sub-category of transcripts which raters can identify readily, because of the difficulty in comprehension, are bizarre attributions of mental states. It is not that the mental state is inherently bizarre, although it might be; but rather that the rater finds it bizarre that anyone should attribute that mental state to a person in that context. Most commonly this occurs in the context of explanations of the subject's own affect. For example: *"I think somewhere ... em ... what makes me feel ... but what makes me feel more rejected is that she breast-fed me. And she didn't breast-feed my sister"*.(-1)

Sometimes such attributions are characteristic of particular psychiatric conditions, although by no means restricted to transcripts from patients with such diagnoses. Bizarre mental state attributions are not restricted to readily recognisable cognitive distortions associated with

psychopathology and may seem bizarre only in the attachment context. For example, Interviewer: ‘Why do you think your parents behaved as they did?’ Interviewee: “*They were over-influenced by the media, particularly television*”. Interviewer: ‘Can you tell me a bit more about this?’ Interviewee: “*The commercial channel just started when I was 4 and ½*”. (0)

Another characteristic of speakers in this group is their focus on the literal meaning of words drawing inappropriate non-reflective conclusions. They often, for example, appear to misinterpret the interviewer by responding to words out of context in the service, it might appear, of rejecting mental state language. This is clearest when the word is actually misunderstood. For example: Interviewer: “Is there a memory that conveys her emotional presence?” Speaker: “*Presents, she always gave us presents!*”

6.4 Disavowal of RF

A very different type of interview, with a little more indication of reflective-functioning comes from individuals who appear to disown their reflective capacity. The absence of mentalizing is comparable to those who reject mentalizing but the repudiation of a reflective stance is accompanied by little or no hostility, indicating that the decoupling of this aspect of thinking does not lead the subject to perceive the task as intrusive. Thus, in place of the aggressive assertion of ignorance of states of mind in self or other, the subject is likely to respond with polite refusal to being quizzed about internal states. Interviewer: “Did you ever feel rejected as a child?” Subject: “*I don’t know. I really could not say*”. (1)

The strength of the denial is equivalent in terms of the sparsity of mental state information but the strategy of implementing this seems to be more passive and evasive. The subject appears to be quite skilled at side-stepping questions referring to their own feelings, reflective capacity or the motivation of the primary caregiver. Most commonly the subject pleads ignorance when directly confronted by questions. Interviewer: “Do you think your childhood experiences have influenced you in any way?” Subject: “*I can’t think of anything ... ehm ... What I learnt from my childhood experiences? Nothing that I can think of at the moment*” It is important to note that replies to questions may seem limited or evasive without the context of the earlier part of the interview. The rater should always consider whether the question has already been answered earlier, e.g., the subject has described in a thoughtful way the emotional impact of having been fostered and sent to boarding-school, but later (in response to the question about the effect of childhood experiences on adult personality), he says that the most important thing was his experiences of being sent away from home. This should be noted by going back to the earlier descriptions and rating from those passages to give the ‘demand question’ rating.

The most common hallmarks of disavowal of RF are concrete explanations of the caregiver’s behaviour in terms avoiding reference to mental states. Explanations may be sociological: “*People in their social class never expected to be affectionate with their children*”(1); referring to external conditions which pertained in their childhood: “*They had no resources to look after us, they sent us to boarding school because it was the best solution to an extremely complex problem*”(1); include generalisations and avoid specificity: Interviewer: “Why do you think your parents behaved as they did?” Subject: “*In all their actions and things they showed caring for*

us; we could not have asked for more, my brother and I had everything which children could possibly wish for.” (1). Reasons for human actions may be found in the physical, rather than the psychological, domain: “My parents were older than those of my friends”, “My mother found it hard when I was young because of her ‘hip thing’”, “We were very close to one another - we lived in the smallest apartment that you have ever seen”, “As a child I was hard to get to know; I had every kind of illness and allergy going; if there was a bug in the neighbourhood I would catch it, I was just that kind of a child, know what I mean”(1, in each case). Explanations for the parents’ behaviour are frequently framed in terms of family structure, such as sibling order, age distribution of siblings or the presence of the parents’ parents: “I just got lost amongst all the children because I was the youngest”, “They didn’t feel they needed to bother because I was the oldest, I was supposed to be able to look after myself” (1). Subjects with this type of transcript may describe their parents’ relationship but this is not linked into their behaviour with the child, is not explained or elaborated: “I think life was hard for them, I mean making a living, raising kids, etc. It was a lot of work, it was a struggle and they didn’t have any inheritance or anything” (1). A common difficulty in rating this aspect of RF is that raters may feel that the explanations offered are in fact plausible, e.g., having a very large family was an important reason for the parents’ behaviour. However, it is not an explanation in mental state terms, and this should be rated low. It may help to remember that reflectiveness is not supposed to be helpful or necessary in all circumstances - higher rating does not necessarily imply greater mental health, common sense or capacity to form good relationships.

6.5 **Distorting or self-serving RF**

In most transcripts there is a degree of self-serving distortion, a bias which reflects the tendency of the mental apparatus to reinforce the cohesion of the self representation. There are at least three powerful biasing forces which may affect what people say concerning the mental states of self and others in relationships.

First, memories of the mental state of self and others may be overwhelmingly egocentric, putting the self in a prominent role, “They thought of little else except what was good for their son, what might please me, they were quite preoccupied by those kind of concerns”(1). Such transcripts are indicative of the subject grossly over-estimating the extent to which the self may have been the cause of the behaviours of others: Interviewer: “What memories do you have about your relationship to your father as anxious?” Subject: “I think from the time I was very small I just made him anxious. He could ... he did not feel comfortable with me. I don’t know why it was but I just provoked this anxiety in him and I think my parents’ marriage failed because of that.”(0). In some transcripts it is evident that the subject erroneously assumes that the self was the intended target of other people’s actions when this is quite unlikely to be the case. A woman with quite rejecting parents recounted the following story. “I remember the time when they were going out for the night and I was as usual very upset that they were leaving me and I cried and cried. They came back quickly. Like in 20 minutes and my mother said the car had broken down up the lane, and they had to abandon it which they did, but I think they were just not wanting to leave me and they went through this whole farce of getting the car mended and everything.” Interviewer: “What makes you think it was because of you and not the car?” Subject: “I just know. I always knew that they sensed what I was feeling”(0)

Second, assumptions concerning the mental state of others are also frequently self-aggrandising; there is a purposeful distortion in the representation of the motives of self and/or others in the direction of enhancing self esteem (e.g. *"I was the apple of my father's eye and my mother was absolutely never jealous about this"*). Distortions are in a specific direction of recalling mental states where they take credit for their success while denying responsibility for their failure. Assumptions about others' negative affects in relation to the self are minimised (*"I was quite a little rascal, but she never felt angry with me"*). When negative feelings and attitudes are remembered, these are likely to be selectively attributed to alternative causes (rather than the self). *For example, a man who was apparently an unwanted child, "an accident" with almost grown up siblings, insisted on attributing his mother's "tiredness" to the "adolescent problems" of his older brother and sister and chose not to consider that he might have also been a part of "the heavy burden that her life had placed upon her"*.

Third, inaccuracies concerning the nature of mental states may often be readily seen as self serving. Some subjects appear to expect wishes to be reality or express the belief that you can read another's mind or that you always know or knew with certainty what another was thinking/feeling. These distortions can give a feeling that the interviewee is totally out of touch with the mental states of those around him but they serve to assert a sense of superiority over others, particularly the parent. Interviewer: "Why do you think your father behaved as he did when you were a child?" Subject: *"By the time I was about eleven, I despised my father for his inadequacy, I thought he was a complete idiot, I knew it by then ..."* (1) Interviewer: Has your relationship to your father changed since childhood? Subject: *"A few years ago, I employed him for awhile, which was a bit stupid, I ended up having to sack him and haven't spoken to him since."* (1)

In summary, the mental state attributions in these transcripts are self-justifying and self-perpetuating. Thus individuals promote the selective availability of mental state information that confirms judgements already arrived at, leading to selective retrieval of memories and over-confidence in such memories. Sometimes the disposition towards self-justification leads to a probable re-writing of memory where, for example, new information which confirms a previously tentatively held view might lead the individual to claim and believe that he had always been confident of that position. Thus new information is sometimes incorporated without the change being registered. All these are self-serving biases, designed to enhance the individual's self-esteem, at the expense of a plausible understanding of oneself or others .

6.6 Naive or simplistic RF

This is an extremely common category of low RF. The hallmark of this type of speech is predictability and the reducibility of mental state attributions to social clichés. The representation of the mental state of the other is one dimensional, it never reflects mixed emotions, conflict or uncertainty about the beliefs and feelings of others. These may be interviews rated secure with little indication of either loss of memory, derogation, ongoing preoccupation or passivity. Nevertheless, it is clear to the rater that the subject's appreciation of intentions of others is only partial either within key paragraphs (instances) or averaging across

the transcript as a whole and they tend to receive a '2'-'3' rating.

The interviews may be marked by 'splitting'. Attachment figures are portrayed as either all good or all bad. (e.g. *"My father was never around, he never bothered and was totally unavailable while my mother was always concerned and there for us"*). A caregiver full of love for the child in an early part of the interview is reported as harbouring malevolent thoughts and feelings. If the subject is aware that this may represent a shift in the attitudes of that caregiver, this is never acknowledged in the narrative.

More commonly the interview is simply very superficial, often totally banal, and the subject makes excessive use of clichés. For example, *"Well because they loved us and wanted to ... give us, a you know, a feeling of security, er, of being wanted um and a, a good start in life I suppose, yeah"* (2). There is little in the interview to 'surprise' the rater and much of what it contains about specific mental states may be derived from the use of the English language - is culturally encoded into language or 'canned'. For example, *"He never took any notice, I don't think he could have cared less about us"*(2). The normalisation of negative experiences in terms of reference to global/cultural trends is a frequent manifestation of the lack of personal discovery which such canned statements reflect; e.g. *"All parents want the best for their children"*(1) or *"All children growing up at that time faced the same type of neglect"*(1). As such statements are rarely specific to personal experience, either of the self or the other, there is a specific overlap between this category and the disavowal category described above. However, whereas disavowal of RF is predominated by concrete explanations in this category mental states are clearly and explicitly referred to even though these are portrayed as being simpler and more uniform than they may be expected to be in reality. For example: *"There were 10 of us children in the family so my mother never had much space in her mind for any of us as an individual"*.(2)

Simplistic responses are not exclusive to this category of low RF interview. Often relatively reflective interviews contain simplistic statements but these are offered in the context of far better elaborated narrative. Evidence of naive simplistic responses should be considered in terms of whether or not this is the best account the speaker can offer, i.e. as a substitute for understanding as opposed to something that accompanies or ends off an otherwise reflective passage.

6.7 Overly-analytical or hyperactive RF

There is a group of interviews which have many of the hallmarks of mentalising yet on closer examination appear to fall way short of the mark. The subject comes across as psychologically-minded but in studying the narrative his/her reflections are mostly irrelevant to the task. The content of the interview, if anything, is excessively deep, with detailed yet unconvincing descriptions of subjective reactions of both subject and others (caregivers, siblings, partners or other attachment figures).

The transcripts reflect a state of affairs where the search for insight is quite compulsive, yet unproductive. Mentalisation spins like a car wheel which has lost contact with ground. Because

so many of the connections between what is thought about thoughts or feelings and the original thoughts or feelings themselves have been lost, metarepresentational capacity works overtime but without real impact. This is slightly different from what is classically talked about as intellectualisation, because intellectualising is generally more coherent, may be more plausible, and does not need to be repetitive and entangled. These subjects may be actually aware of failing to arrive at meaning, as the example below illustrates. However, they are drawn to this way of talking in the hope of re-engaging their freewheeling thinking with a sense of affective significance. But being drawn, with the use of prompts, into focusing on the content of the material they present gets nowhere, leading to only further elaborations without the feeling of real contact or the possibility of movement.

An individual was sent to a series of boarding schools and therefore frequently separated from his family home; he discovered at some stage in his childhood that his father was not his biological father. His parents had violent quarrels and finally divorced:-

Interviewer: "Would you say that your childhood has had a big influence on who you are today?" Subject: *"Well, it's difficult to answer that because, you know, raised in the socio-cultural context that one is, er, you know, you take it in almost as soon as you start to read, that the child is father to the man, all that ... What's that? Wordsworth, isn't it?" Is there anything in particular? "I mean, again I suppose logically, I mean ... You don't have to put it all down to post psychoanalytic cultures and all that ... There seems to be a certain inevitable logic to the fact that you go through a necessary formative phase, well phases, presumably they are bound to affect your subsequent development."* (1) Interviewer: "Was there anything that you would consider a setback in your development?" Subject: (sighs) *"Well, interesting, you should ask that ... I mean ever since the concept of maturity, you know, formed in my consciousness I have always been aware of being some distance from it. It's so easy to rationalise, you come out with trite, er, er, half-baked, er, psychological interpretations."* (1)

Explanations are frequently characterised by excessive length. When giving accounts in mental state terms subjects go on apparently interminably. This might be in contrast with clear and coherent 'factual' accounts they are able to give of historical events, physical circumstances, or sometimes even the mental states underlying their own or others' behaviour when the relationships are distant, non-attachment relations. As soon as the narrative turns to attachment relationships, the depth of their narrative goes well beyond what could be appropriate to the interview situation, and frequently hard to follow. The text can come to resemble a travesty of a psychotherapy session with the subject failing to show convincing understanding of their internal world. For example, one subject intending to give an account of her growing understanding of the exploitative relationship between her mother and father said: *"I began to see that it takes two to tango ... It was a perfect collusion between the two of them. What has been called in popular psychology, you know, the doormat-tyrant relationship ..."* The subject offers too many interpretations without adequate support, and may fail to filter out contradictory or oscillating stances.

There is a clear assumption on the part of most of these subjects that their concerns, language and understanding are shared with interviewer giving a presumptuous feel to the interview. It is

as if they assumed that as the interviewer is a ‘psychological/ mental health type’, their extreme analytic stance is justified. The interviewee clearly loses sight of what the interviewer may be perceiving or thinking and is not fully engaged with the interviewer. For example: *“You will be interested in this. My mother I think was a very passive woman, who was so dependent on my father because of her self-perceived inadequacies that she was terrified of asserting her rights to equal treatment. Of course she got depressed. But remember, this was before the feminist revolution. She had no one to turn to except her children who of course failed to understand what she needed. I believe this happened to a number of women at that time - they drew on the emotional resources of their child to support their weakened sense of self. A kind of narcissistic personality distortion, I think you would call it. I actually don’t think that labelling helps very much in these situations. My father suffered of the same.”*(2)

In the end the interviews reflect diffuse mental states where the mental state of the protagonists are not at all clearly marked out. Based on extracts it may be hard to distinguish such interviews from reflective ones. Taken as a whole interview these transcripts stand out because of the confusing nature of the narrative, the dominance of inappropriate jargon, the odd use of mental state terminology, and the ultimate shallowness of the insights offered. The way we understand such productions is that features of meta-representations are present without connection to the representations of events which have formed the basis for a generalisation. For example: *“Well I think that’s probably ... aware of my anger when he was around and that if rejection is a lack of love, the inability of people to love and therefore just like cutting off, I was certainly in a constant state of rejection.”* Here, the phrases ‘lack of love’, ‘inability to love’, ‘cutting off’, ‘state of rejection’ are used without a sense that the actual experiences which entailed the pain of abandonment to which these phrases ostensibly refer are currently available to the subject. It is essential to remember that if a plausible context or elaboration has been provided earlier in the interview, the experiences should be considered to have been described in the rated passage.

7 RATING PASSAGES

7.1 Rules for identifying passages: Demand vs. permit questions

Passages are to be identified according to the context in which they appear, specifically the question posed by the interviewer and the explanations and previous information the speaker has provided. Questions in the interview transcript may be divided into two types: (1) those that permit the speaker to demonstrate their reflective-self capacities (e.g. orient as to background; what did you do when you were upset as a child?); and (2) those that demand from speakers a demonstration of their capacity for reflective-self function. An exhaustive list of the latter questions in the AAI follows:

-why did your parents behave as they did during your childhood?

-do you think your childhood experiences have an influence on who you are today?

-any setbacks?

-did you ever feel rejected as a child?

-in relation to losses, abuse or other trauma, how did you feel at the time and how have your feelings changed over time?

-have there been changes in your relationship with your parents since childhood?

When applying the principles in this manual to other research interviews, raters should identify a set of questions which are most likely to elicit explanations in terms of mental states. Passages in response to these demand questions must be rated, and should later be taken into account when arriving at a global rating of the interview. In some interviews, interviewers might use prompts which in effect are demand questions (e.g. ‘And why do you think they did that?’) and passages which follow such prompts should be treated in the same way as passages following the questions listed above. If the speaker has already addressed a demand question in their response to a previous question, no penalty is incurred. Non-reflective responses to permit questions, i.e. all other questions, should not carry as much weight as they would if provided in response to a demand question. When combining the scores, highly rated responses to these other questions should however contribute to the overall rating, and it should be noted when subjects are highly reflective in response to a non-demand question, which suggests a habit of spontaneously thinking in terms of psychological explanations. An answer to each question will generally be rated as a single passage, unless it contains more than one ‘answer’ within it, e.g., “Why did your parents behave as they did?” might be answered with a clearly different way of thinking for mother and for father. The demand question about loss should always be rated separately for each loss.

7.2 Guidelines for rating identified passages:

Six scale points are defined below. Based on reviewing the definitions of RF given above, raters should assign a score to all responses to demand questions, as well as any other responses that merit attention. Some effort should be given to assigning a typology to each low rating in order to facilitate final classification.

-1 or Negative RF

Passages which are rated as negative or ‘0’ must be distinctly anti-reflective or bizarre/inappropriate.

An anti-reflective passage expresses hostility or active evasion in response to an opportunity for reflection. Sometimes the subject may express overt hostility by criticising the interviewer or the task. More commonly, however, active evasion consists of trying to distract the interviewer from the task by, for example, starting a conversation on an irrelevant topic or disengaging from the task by engaging in any activity which precludes complying with the demands of the task.

Sometimes these reactions are non-verbal, for example: getting up to make a telephone call or going totally silent.

Bizarre explanations of behaviour unequivocally invoke mental states in self or other which are beyond the bounds of common-sense psychology or even poorly-applied theory-driven insight. To be rated negative the passage must be impossible to understand without making the assumption of 'irrationality' on the part of the interviewee. Bizarreness usually arises contextually, in other words in relation to other statements which the subject has made. Complete non-sequiturs, over-familiarity, gross assumptions about the interviewer on the part of the subject are examples. These are also contextually bizarre except the context in this case is the interview situation as a whole rather than a specific part of the interview.

1 Absent but not repudiated RF

This rating applies to passages in response to demand questions, or whenever (in the rater's view) there is an implied demand for a mentalising response to a probe.

This rating is given when the subject does not mention mental states, in spite of a clear opportunity to do so. Passages rated as '1' may be sociological, excessively generalised, concrete or overwhelmingly egocentric. Passages can only be rated as '1' if the rater is no better off in terms of knowledge of the mental state of subject, caregiver or other, having read the passage, than he/she was before reading it. Even if the rater suspects the accuracy of the subject's representation of mental states, this would be insufficient grounds for rating the passage as '1' unless the passage is grossly self-serving. Thus it is not sufficient to be inaccurate to get a '1' rating, the passage has ego-centric and indicate that the interviewee is quite out of touch with the mental states of those around him/her.

3 Questionable or low RF

This rating is given to the quite common passages where the rater is uncertain if the passage represents genuine reflective-functioning or just a 'canned' statement which is produced in response to the interviewer's demand but is not underpinned by genuine reflective-functioning. The passage may make use of mental state language but there is an absence of material which would support the assumption that the subject genuinely understood the implications of their statement. In general this rating would be given when the passage contains no evidence of: a) awareness of the nature of mental states, b) an explicit effort to tease out mental states underlying behaviour, c) a recognition of the developmental aspects of mental states or d) interaction with the interviewer indicative of the awareness of the mental states of that person. This is a borderline category. The rater should focus on whether the statement is 'obvious' and could be said simply as a 'manner of speaking'. If a statement is counter-intuitive in that context even though it is 'canned' a higher rating may be appropriate.

Most frequently '3' statements are somewhat banal and superficial. The reflective-functioning may be there by suggestion but it is unclear and is as likely to be a cliché as a proper reflective statement. Most of the statements could be seen as reflective if the rater were to give the subject

the ‘benefit of the doubt’. Also, statements which are definitely mentalising but excessive and irrelevant to the task (overanalytic) may be given a ‘3’ rating. These statements are considered to be on the borderline because although they are psychologically minded they do not carry conviction. Thus again the rater may be uncertain if the statement should be credited or not.

If the phrase appears to fit the descriptions for moderate-to-high RF statements given in section 3 of the Manual, a score of ‘3’ would be inappropriate. The rater should have a reason why a statement of that sort cannot be confidently assigned a higher rating. Other than the ‘canned’ nature of the reply, common reasons might be strong prompting from the interviewer, a failure to be fully explicit, and the rater’s uncertainty if he is not attributing mental state reasoning to the subject on the basis of his/her own understanding of the situation described.

5 Definite or ordinary RF

Passages rated ‘5’ must have a feature which makes reflection explicit. This is the most important criterion. If there is no explicit reference to either the nature of mental states, how mental states relate to behaviour, the properties of mental states or mental states in relation to the interviewer, then the rater should consider assigning a lower mark to the passage. Even if the mental state is fairly simple, if it is described clearly and briefly reflected on in a way which does not suggest a restatement of what might be socially expected, this rating may be appropriate

Other than being explicit, the statement does not need to reflect sophistication. Although the statement should not be a cliché, it may be fairly ordinary, not reflecting particular insights or sensitivities. Normally the passage fits fairly well under one of the categories listed under the examples given in section 4 above. If the passage contains a number of the features listed in that part of the manual, the rater should give the passage a higher rating, unless it is also flawed in an important way (e.g., overanalytical style). One of the six reasons, below, for rating ‘7’ may be present in the context of a very simple observation of mental states which would otherwise rate only a ‘3’; in these cases, a ‘5’ is likely to be appropriate.

7 Marked RF

Passages rated ‘7’ are usually broader than those rated ‘5’ but essentially they meet similar criteria. These passages are rated higher for one of six reasons.

1) The passage may contain a sophisticated statement concerning mental states which fits the description of at least two of the categories listed in the section on indications of moderate to high RF. Here the rater is looking for passages readily classifiable as being reflective. The statement owes its ‘obviously’ reflective nature to combining several features of mentalising such as awareness of the limitations of knowledge of another’s mental states, as well as indicating a recognition that individual perspectives on the same objective event may differ.

(2) The passage may be ‘marked’ in reflective-functioning because the view of mental states presented by the subject is unusual and surprising to the rater. Passages which cast an original

perspective, which nevertheless is readily understandable to the rater, reflect mentalization on the part of the subject. Raters should however be aware of the possibility of “borrowed” reflective-functioning, where the subject is repeating ideas presented to him/her in other contexts (family legends, therapeutic consultations, etc.). In such instances a rating of ‘3’ would probably be more appropriate.

(3) The passage may be complex or elaborate in that the mental state of the self or the other is described in unusual detail. Raters should look out for the presentation of complex, multi-layered mental states, conflicts, mixed emotions, false beliefs and the like. It is important that for a rating of ‘7’ the passage should indicate that the multiple mental states attributed to the person are considered alongside one-an-other and in relation to one another.

(4) A rating of at least ‘7’ (but more likely ‘8’ or ‘9’) should be given to passages where mental states are placed within a causal sequence. By this we mean that the subject considers (a) how the mental states arose (what perceptions of reality lead to the belief or desire assumed), (b) how the mental state influenced behaviour and (c) what impact or implication the mental state has subsequent perceptions, beliefs and desires.

(5) A somewhat overlapping but occasionally separate reason why a “marked” rating is appropriate rests in the interactional perspective on mental states manifested by some subjects. Interactions refer to seeing mental states as impacting on one another in a causal way. Such interactions may be of two kinds. (a) More commonly subjects explicitly state how the mental state of one person may impact on the mental state of another. This may be in the context of the feelings of one affecting the feeling state of another, or how perceptions of the expression of feeling in one person are construed by another, or how experiences of certain circumstances are interpreted by one person differently from another and how these different perceptions affect respective behaviours etc. (b) Less frequently subjects may consider the interactions of mental states within a single mind. Here the common examples involve conflicting perceptions or desires and mixed emotions and the subject conveys some reconciliation of these. If the process of integration is described with appropriate elaboration or illustration, the rater should consider awarding a higher rating to the passage (‘9’).

(6) If a subject acknowledges a particularly difficult situation, with the thoughts or feelings appropriate to that, then credit is given for the subject’s willingness to accept experiences rather than defend against them, avoid rationalising the behaviour of people who hurt him or her, etc..

Illustration: Interviewer: “were either of your parents ever threatening to you in any way, either for discipline or jokingly?” Subject: *“My father was threatening to me once ... I was being a pain, I wouldn’t do something, my mum and I had an argument about it. She was very upset and I was being stupid. I wouldn’t acknowledge that I was upsetting her so much. My Dad found me and came and shook me, and told me how much I had hurt my mother. He was quite rough and I was really scared. He came back later and said he was sorry and explained, and we made it up.* Interviewer: “What did you feel about that?” Subject: *“I can remember thinking “Ah, I’ve got him! ... I shouldn’t have felt it really, but there was a feeling of triumph.”* (7)

9 Full or Exceptional RF

The difference between this rating and the previous rating is as much one of degree or quantity as of quality. The same reasons which may lead a rater to assign a “marked” rating may lead him/her to consider awarding an “exceptional” score. The difference lies in the amount of sophistication shown, the degree of complexity presented, the completeness of the causal account, the degree of ‘surprise’ the rater experiences at the subject’s understanding, the intricacy of the interaction between mental states offered etc. For a rating of ‘9’ the passage must be unusual in at least one of these respects.

A further circumstance which might justify the award of this rating even to passages which would normally be rated “marked” is the context in which the passage appears. If the rater sees the part of the narrative as particularly emotionally charged and difficult for the subject then showing even marked levels of mental state understanding may be considered “exceptional”. Examples might include the understanding of rejection, neglect or abuse by the caregiver in childhood, or understanding feelings of current anger or resentment from or toward the attachment figure. The rater should note that the presence of mentalising may give the impression that the experience recounted was not exceptionally difficult. To circumvent this the rater should take an “objective” (almost sociological) view of the difficulty involved in the experience for that person and adjust ratings accordingly. For example rejection by one parent may not be quite so difficult for someone who appears to be assured of the love and commitment of the other parental figure. By contrast, an individual whose history reveals no such ameliorating factor might be more readily credited with an “exceptional” rating if the understanding they show of the experience of rejection meets the criteria for “marked” RF, even if the understanding does not come across as hard-won.

The most common justification for an “exceptional” rating is the apparent awareness of the subject of important aspects of the mental states of all protagonists within an interaction, where the protagonists are placed in relation to one another in terms of their feelings and beliefs and these are sufficiently complex and elaborate for the rater to be persuaded of their accuracy. The subject offering such a “full” picture may not be exceptionally insightful, although the passage must contain sufficient “surprise” and coherence for the rater to feel it is unlikely to have come from contaminating sources (e.g. regurgitation of a therapist’s or parent’s view). The passage should therefore have a personal character, i.e., experienced as personally significant and meaningful, and may seem to be developing further during the interview itself.

In an ordinary sample we would expect less than 10% of scored passages to receive such a high rating across interviews.

7.3 Rules for aggregating reflective-functioning ratings into a single score for Adult Attachment Interview

Although each identified passage of the interview, particularly the responses to demand questions, is scored by raters, the almost infinite number of permutations which may end up characterising an interview means that there is no formula for arriving at a global rating. Rather,

the rater has to consider the interview as a whole, alongside the ratings for individual passages. The rater should not take an arithmetic average of the ratings given to core passages or even to all passages in the transcript. Too little is known of the psychometric properties of the individual ratings to permit this simple expedient. The rater has to come to a judgement about the whole text on the basis of this manual, and over time in relation to their experience of other narratives to which they have assigned ratings.

Although some simple rules for assigning a rating to the interview are available, none of these should be used blindly and they should not override the rater's impression of the entire interview. For example an interview may contain striking examples of negative RF at the same time as containing one or two examples of marked RF. In such a case, none of the formulae offered for assignment of a global rating apply and the rater has to decide if the best fitting category is '1' or '3'. The rater will have to make a judgement if RF may be considered to be absent in an attachment context ('1') or is relatively low and patchy ('3').

In making higher ratings the rater should also consider the total interview. For example, quite commonly a particular passage in the interview may not merit a '5' rating, but when considered in the context of the whole narrative the same text may clearly merit a higher rating. It is vital to try to avoid a halo effect, so that interviewees who are likeable or easy to sympathise with are not rated higher, and not to rate somebody lower because they or their attitudes are unappealing. If their statements are reflective, they must be credited. To take a common example, a person who is egocentric or self-serving should be rated as reflective if they do make statements meeting criteria in the manual, *unless* the perceptions and explanations *in themselves* are self-serving. Sometimes the rater may not be able to decide between adjacent categories. For this reason, classifications are anchored at odd numbers. When the rater is confident that a particular transcript falls in between two classes, he/she should assign the even number inbetween the classes to the transcript.

7.4 The overall interview rating

-1 Negative RF

This overall rating should only be given to interviews, very rare in normal samples, where the interviewee systematically resists taking a reflective stance throughout the interview. The subject is either hostile to the notion of reflection expressed in derogation and dismissal of any attempts on the part of the interviewer to initiate such reflection or is so confused about their attempts at reflection that the rater may be said to be almost "shocked" by the utterances. In either case, for a rating below '1' to be given, the rater should be certain that not even a single reflective item can be marked in the interview as '5' or above. In interviews where either a rejecting or a bizarre stance is observed alongside some rateable RF passages, ratings between '1' and '3' should be considered depending on the balance of items found.

Common Types

-1(A): Rejection of RF

Interviews of this type rated ‘-1’, respond with hostile refusal to at least three core RF questions. In addition, there are some general characteristics of the interview such as a lack of participation in the interview process, overt hostility to the interviewer, evasiveness and marked incongruences. If the interview fits the latter descriptors but has less than three but at least one striking hostile refusal, and no clear instance of RF (5 or above) is found, the interview should be rated ‘0’.

-1(B): Unintegrated, Bizarre or Inappropriate

This is a rare category and a literally puzzling one for the rater. Its hallmark is that mental state attributions are hard to understand. To place an interview in this category the interview should contain at least three examples where an inexplicable, bizarre or inappropriate attribution was made by the subject. These need not occur uniquely in response to core RF questions. It is, however, insufficient for the answer to be unusual or simply odd. The rater’s reaction is likely to be one of shock that anybody could make such an attribution in such a context. As an extreme example, frankly paranoid responses or thought disordered ones would create this kind of subjective reaction. In addition there are some general features of such interviews including a lack of meaning, a lack of explanation, a comprehensive avoidance. If the rater identifies at least three instances of bizarre explanations of behaviour, either paranoid, thought disordered or highly incoherent and therefore impossible to understand but not thought disordered, then the interview should be rated -1. If the interview is generally poorly integrated or somewhat bizarre in terms of mental state attributions but there are less than three clear passages of this type and no passages are rated ‘5’ or above, the interview should be assigned ‘0’.

1 Lacking in RF

This rating should be given to interviews where reflective-functioning is totally or almost totally absent. In these interviews, mental states may occasionally be mentioned with regard to the self or other, but these give no picture of the subject’s or the caregiver’s beliefs and feelings underlying behaviour. Usually, no mentalizing statements are present. Alternatively, such statements are clearly inaccurate and full of misunderstanding and contradiction. In all cases, mentalisation is absent in the narrative and awareness of the nature of mental states, if present, is only discernible by inference.

Common Types

1(A) Disavowal

There are at least three instances in the transcript of assertion of ignorance concerning mental states. Alternatively, there are comparable examples in these core passages of evasion of questions, physicalistic, behavioural or sociological accounts and global and generalised statements concerning the behaviour of the other or the self. In general terms, such accounts tend to be barren, lacking in mentalizing detail and mentalizing phrases are restricted to those in use in ‘common parlance’ (canned explanations). A certain concreteness tends also to characterise such interviews. In order to assign this category there can be no instance of RF rated above ‘3’.

1(B) Distorting/self-serving

These interviews do contain reflection but are seen by the rater as flawed. Principally reflective passages and answers to core questions are egocentric, exaggerate the importance of the narrator, are overly favourable to the subject or are self-serving to the point where the accuracy of the representation of the mental state of the other may be justifiably called into question. A key bias to the depiction of mental states is social desirability – the subject wishes to present himself in a favourable light. These distortions can lead to marked inconsistencies in the presentation of the mental world of both self and other. Subjectively, the rater is most likely to feel a strong sense of irritation with such interviews. To use this category the rater should note at least three instances in the transcript of such purposeful distortions in response to core RF questions. Further there should be no instance of RF rated above ‘5’, or more than two above ‘3’.

3 Questionable or Low RF

This rating is given to transcripts which contain some evidence of consideration of mental states throughout the interview albeit at a fairly rudimentary level. Interviews receiving a ‘3’ will contain elements of a reflective stance, for example the interviewee may consider developmental or intergenerational elements which are not seen by the rater as banal (i.e. a lower rating) but are nevertheless not specific enough to the individual's personal experience to merit a higher rating. An interview rated a ‘3’ may contain more than one example of reflective-functioning at level ‘5’ or above but no rating as high as ‘7’. Further, it must contain at least three instances of a ‘3’ or ‘4’ rating. In reading many of these transcripts the rater may intuitively wish to attribute to the interviewee a relatively good capacity to reflect (‘5’ overall) but, in the absence of concrete evidence for this in terms of at least three ratings at ‘5’ or above, this may be indicative of the rater’s rather than the interviewee’s capacity. For the most part, references to mental states and their impact on behaviour are not elaborated. Also, a number of relatively reflective passages may be counterbalanced by negative ratings elsewhere (although not of sufficient frequency to warrant a lower rating). However, if the rater, through taking the interview as a whole, is readily able to bring together aspects of the interview which would clarify the mentalising content of several passages, a higher rating should be given to the interview.

Common types

3(A) Naive-simplistic

These transcripts show a partial appreciation of intentions of others either within instances or across transcript as a whole. The understanding offered may be very superficial, often totally banal, with excessive use of clichés in referring to mental states and perhaps normalisation of negative experiences beyond what may be culturally accepted. The understanding is rarely specific to personal experience, either of self or other, does not enter into the complexities of mental states (conflict, ambivalence etc) . The interview is likely to contain ‘canned’ statements and the interviewer is likely to experience a sense of superficiality about such transcripts. To assign this category the rater is expected to have identified naive simplistic passages as the majority of low ratings, and no ratings as high as ‘7’.

3(B) Over-analytical or hyperactive RF

This is an important but somewhat difficult category. To the naive rater such transcripts may seem quite reflective. In fact one of the hallmarks of the category is that the interview has somewhat greater depth than might be expected. The interview is diffuse, however, and the insights offered are unintegrated and do not link in a compelling way to the individual's experience. The interview is given this rating if there are at least three instances in which the speaker is overly-analytical; if one or more of these includes text that qualifies as bizarre reasoning, distorting/self-serving, the rater should assign the interview as a whole a '1' or '2' rating. The overly-analytical passage/interview may also be rated '1' or '2' rather than '3' if there is little or no sense over the course of the interview of a developing understanding. If the rater is considering assigning an interview to this category but the number of 'definite' or 'marked' RF instances make this assignment questionable, the rater should take a further look at these in order to ascertain the appropriateness of the original rating.

3(C) Miscellaneous low RF

There are transcripts, relatively few in number, where a '3' rating applies even though the transcript is neither particularly naive nor overly analytic. Most commonly the rating is arrived at as a compromise between a higher and a lower rating. For example, transcripts which show clear evidence of disavowal, yet contain definite evidence of reflective-functioning should be assigned to this borderline category. Other transcripts have numerous RF passages but none of them go beyond the '5' - '6' range, in addition to one or two '1s' in response to demand questions. Yet others may meet the requirement for a '5' rating but taken as a whole the transcript is not really reflective. In these transcripts the speaker's model of mental states has to be partly inferred by the rater (e.g. emotional events are vividly described, but without the implications being spelt out, there may be little awareness of a link between cognition and affect, the impact of one relationship on another, and so on).

5 Ordinary RF

This is the most common rating in a high functioning "normal" sample. Transcripts at this level should have a number of instances of reflective-functioning even if these are prompted by the interviewer rather than emerging spontaneously from the interviewee. In contrast to interviews rated '0'-'4', interviews rated '5' give convincing indications to the rater that speakers have some kind of a model of the mind of attachment figures as well as a model of their own mind which is relatively coherent even if it is simple, and is unlikely to have been solely derived from shared culture rather than from personal experience. Whereas in interviews given ratings of '4' there may be a model of the mind akin to models observable in interviews receiving high ratings, only in these latter cases is the model easily understood and well-integrated.

There may be transcripts where the subject's mentalising stance is attenuated by difficulties in expression, but nevertheless the rater should exercise discretion and generosity in giving the '5' rating if some clear instances of mentalising are present of at least a '5' level. It is not expected that such interviews should contain passages which would merit a 'negative' or a '1' rating without an expansion immediately following or elsewhere in the transcript which indicates some

reflection on the same subject. However, ratings of '5' would most commonly be given to interviews which combine statements that are genuinely reflective ('7') with rudimentary, more superficial ones ('3').

Common types

5(A) Ordinary understanding

The subject shows a capacity to make sense of their experience in terms of thoughts and feelings, and has a consistent model for this which needs little or no inference from the rater. However, the model is limited, and would not tackle the more complex aspects of interpersonal relationships, e.g. conflict or ambivalence, which are not necessarily well understood. The transcripts need to have at least three instances warranting a '5' rating, and no breakthroughs of rejection or bizarre explanations, pervasive disavowal, etc..

5(B) Inconsistent level of understanding

The subject appears to be achieving a higher level of understanding in some parts of the interview, so that certain passages may achieve scores of '7' or even higher. However, the understanding cannot be maintained in relation to more problematic areas of the subject's interview, such as a conflictual relationship with one parent, or losses. These parts of the interview would nevertheless not be expected to fall below a rating of '1'-'2'.

7 Marked RF

These interviews have numerous statements indicating full RF, which show awareness of the nature of mental states, and explicit attempts at teasing out mental states underlying behaviour. Normally, awareness of mental states is clear throughout the interview with frequent passages where the speaker has arrived at an integration of the states of minds of those around them or their own state of mind which the rater may find surprising in the sense of not having thought of it himself or herself. There is also much detail about the thoughts and feelings of all the protagonists and the implications of mental states are regularly spelt out. The person is usually also able to maintain a developmental (intergenerational) perspective. As a whole, the interview gives the rater the feeling that the speaker has a stable psychological model of the mind of his caregivers and his/her own reactions to their mental states.

Any single passage may only illustrate one of the features of full RF listed above, but the interview as a whole gives the impression of someone who is applying a reflective stance fairly consistently to at least one context in their life, e.g. relationship with mother, or less consistently to a number of contexts. Across the interview, one would expect a number of different types of examples of RF but the reader can identify limitations in terms of type or application to context. In order to assign a global rating of '7' the rater should identify at least three passages which may be given this rating or higher anywhere in the interview, no passages which obtain a rating of '1' or lower (without reason; i.e., this does not apply if for instance there seems to have been no rejection to be discussed, and the only loss was of somebody not close, or during very early childhood) and no more than three passages where the rating is less than a '5' in response to core questions.

9 Exceptional RF

These transcripts are rare. They show exceptional sophistication, are surprising, quite complex or elaborate and consistently manifest reasoning in a causal way using mental states. A '9' rating would be given to transcripts which show a consistent reflective stance across most important contexts. A '9' rating for a single passage is given where several aspects of RF are integrated into a unified, fresh perspective. Where three such passages are noted in any single interview, the rater should assign a '9' rating to the interview as a whole. Across the interview, many aspects of full spontaneous RF would be shown in the discussion of different relationships at different times in the speaker's life history. It is unlikely that such an interview would have many passages rated '3' and many would be rated '5' or '7'. If the transcript does not meet the criteria above, yet the rater 'feels' the transcript to be exceptional a rating of '8' should be considered. If only a single '9' rating is present and/or there are more than a couple of examples of questionable RF (3) a rating of '8' is likely to be too high and '7' should be considered.

Acknowledgements

We are very grateful to Abby Schachter for help in preparing the previous version (3.3) of this manual. We also thank Dr Arietta Slade who carefully read, commented upon and helped prepare this version of the manual. The appended scoring outline has been modified from one provided by Cassandra Cook, PhD, for version 4.1, and is intended as a quick reference for raters. This version of the manual has been improved through discussion with groups using it, especially those led by Dr Susan Coates, Dr Arietta Slade and Dr Imre Szecsödy.

References

- Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). Patterns of attachment: A psychological study of the Strange Situation. Hillsdale, NJ: Erlbaum.
- Auerbach, J. S. (1993). The origins of narcissism and narcissistic personality disorder: A theoretical and empirical reformulation. In J. M. Masling & R. F. Bornstein (Eds.), Psychoanalytic Perspectives on Psychopathology (pp. 43-110). Washington, D.C.: American Psychological Association.
- Auerbach, J. S., & Blatt, S. J. (1996). Self-representation in severe psychopathology: The role of reflexive self-awareness. Psychoanalytic psychology, 13, 297-341.
- Baron-Cohen, S. (1995). Mindblindness: An Essay on Autism and Theory of Mind. Cambridge, MA: Bradford, MIT Press.
- Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (1993). Understanding Other Minds: Perspectives from Autism. Oxford: Oxford University Press.
- Bidell, T. R., & Fischer, K. W. (1994). Developmental transitions in children's early on-line planning. In M. M. Haith, J. B. Benson, R. J. Roberts, & B. F. Pennington (Eds.), Development of future-oriented processes. Chicago: University of Chicago Press.
- Bion, W. R. (1962a). Learning from Experience. London: Heinemann.
- Bion, W. R. (1962b). A theory of thinking. International Journal of Psycho-Analysis, 43, 306-310.
- Bolton, D., & Hill, J. (1996). Mind, meaning and mental disorder. Oxford: Oxford University Press.
- Bretherton, I., Fritz, J., Zahn-Waxler, C., & Ridgeway, D. (1986). Learning to talk about emotions: A functionalist perspective. Child Development, 57, 529-548.
- Busch, F. (1995). Do actions speak louder than words? A query into an enigma in analytic theory and technique. Journal of the American Psychoanalytic Association, 43, 61-82.
- Cassam, Q. (Ed.). (1994). Self-knowledge. Oxford: Oxford University Press.
- Churchland, P. S. (1986). Neurophilosophy. Cambridge, Mass: MIT Press.
- Cicchetti, D., & Beeghly, M. (1987). Symbolic development in maltreated youngsters: An organizational perspective. In D. Cicchetti & M. Beeghly (Eds.), Atypical symbolic development. New Directions for Child Development (Vol. 36, pp. 5-29). San Francisco: Jossey-Bass.
- Davidson, D. (1980). Agency. In D. Davidson (Ed.), Essays on actions and events (pp. 43-61). Oxford: Clarendon Press.
- Dennett, D. (1987). The intentional stance. Cambridge, Mass: MIT Press.
- Dennett, D. (1988). Precis of 'The intentional stance' with peer commentary. The Behavioral and Brain Sciences, 11, 495-546.
- Dennett, D. C. (1978). Brainstorms: Philosophical Essays on Mind and Psychology. Montgometry, VT: Bradford.
- Dunn, J. (1996). The Emanuel Miller Memorial Lecture 1995. Children's relationships: Bridging the divide between cognitive and social development. Journal of Child Psychology and Psychiatry, 37, 507-518.
- Emde, R., Kubicek, L., & Oppenheim, D. (1997). Imaginative reality observed during early language development. International Journal of Psycho-Analysis, 78(1), 115-133.

Epstein, S. (1983). Mother-Father-Peer Scale. Unpublished manuscript, University of Massachusetts, Amherst.

Eysenck, H., & Eysenck, F. (1975). Manual of the Eysenck Personality Questionnaire. London: Hodder and Stoughton.

Fairbairn, W. R. D. (1952). An Object-Relations Theory of the Personality. New York: Basic Books.

Fischer, K. W., & Ayoub, C. (1994). Affective splitting and dissociation in normal and maltreated children: Developmental pathways for self in relationships. In D. Cicchetti & T. H. Toth (Eds.), Rochester Symposium on Developmental Psychopathology: Vol. 5. Disorders and dysfunctions of the self (pp. 149-222). Rochester, NY: University of Rochester Press.

Fischer, K. W., & Farrar, M. J. (1987). Generalisations about generalisation: How a theory of skill development explains both generality and specificity. International Journal of Psychology, 22, 643-677.

Fischer, K. W., Kenny, S. L., & Pipp, S. L. (1990). How cognitive processes and environmental conditions organize discontinuities in the development of abstractions. In C. N. Alexander, E. J. Langer, & R. M. Oetzel (Eds.), Higher stages of development (pp. 162-187). New York: Oxford University Press.

Fischer, K. W., Knight, C. C., & Van Parys, M. (1993). Analyzing diversity in developmental pathways: Methods and concepts. In W. Edelstein & R. Case (Eds.), Constructivist approaches to development. Contributions to Human Development (Vol. 23, pp. 33-56). Basel, Switzerland: S. Karger.

Fischer, K. W., & Pipp, S. L. (1984). Development of the structures of unconscious thought. In K. Bowers & D. Meichenbaum (Eds.), The unconscious reconsidered (pp. 88-148). New York: Wiley.

Flavell, J. (1982). On cognitive development. Child Development, 53, 1-10.

Fodor, J. (1987). Psychosemantics. Cambridge, Mass: MIT Press.

Fonagy, P. (1989). On tolerating mental states: Theory of mind in borderline patients. Bulletin of the Anna Freud Centre, 12, 91-115.

Fonagy, P. (1991). Thinking about thinking: Some clinical and theoretical considerations in the treatment of a borderline patient. International Journal of Psycho-Analysis, 72, 1-18.

Fonagy, P. (1997a). Attachment and theory of mind: Overlapping constructs? Association for Child Psychology and Psychiatry Occasional Papers, 14, 31-40.

Fonagy, P. (1997b). Can we use observations of infant-caregiver interactions as the basis for a model of the representational world. Psychoanalytic Dialogues, 7, 207-213.

Fonagy, P., & Higgitt, A. (1989). A developmental perspective on borderline personality disorder. Revue Internationale de Psychopathologie, 1, 125-159.

Fonagy, P., Leigh, T., Steele, M., Steele, H., Kennedy, R., Mattoon, G., Target, M., & Gerber, A. (1996). The relation of attachment status, psychiatric classification, and response to psychotherapy. Journal of Consulting and Clinical Psychology, 64, 22-31.

Fonagy, P., Steele, H., & Steele, M. (1991a). Maternal representations of attachment during pregnancy predict the organization of infant-mother attachment at one year of age. Child Development, 62, 891-905.

Fonagy, P., Steele, M., Moran, G. S., Steele, H., & Higgitt, A. C. (1991b). Measuring the ghost in the nursery: a summary of the main findings of the Anna Freud Centre/University College London parent-child study. Bulletin of the Anna Freud Centre, 14, 115-131.

Fonagy, P., Steele, M., Steele, H., Higgitt, A., & Target, M. (1994). The Emmanuel Miller Memorial Lecture 1992. The theory and practice of resilience. Journal of Child Psychology and Psychiatry and Allied Disciplines, 35, 231-257.

Fonagy, P., & Target, M. (1995). Understanding the violent patient. International Journal of Psycho-Analysis, 76, 487-502.

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

Freud, S. (1911). Formulations on the two principles of mental functioning. In J.Strachey (Ed.), The standard edition of the complete psychological works of Sigmund Freud (pp. 67-104). London: Hogarth Press.

Freud, S. (1914). On narcissism: An introduction. In J.Strachey (Ed.), The standard edition of the complete psychological works of Sigmund Freud (pp. 67-104). London: Hogarth Press.

Frosch, A. (1995). The preconceptual organization of emotion. Journal of the American Psychoanalytic Association, 43, 423-447.

Gergely, G., & Watson, J. (1996). The social biofeedback model of parental affect-mirroring. International Journal of Psycho-Analysis, 77, 1181-1212.

Green, A. (1975). The analyst, symbolization, and absence in the analytic setting. International Journal of Psychoanalysis, 56, 1-22.

Grice, H. P. (1975). Logic and conversation. In R.Cole & J.Morgan (Eds.), Syntax and Semantics: Speech Acts (pp. 41-58). New York & Academic Press.

Harris, P. L., Johnson, C. N., Hutton, C., Andrews, G., & Cooke, T. (1989). Young children's theory of mind and emotion. Cognition and Emotion, 3, 379-400.

Hopkins, J. (1992). Psychoanalysis, interpretation, and science. In J. Hopkins & A. Saville (Eds.), Psychoanalysis, mind and art: Perspectives on Richard Wollheim (pp. 3-34). Oxford: Blackwell.

Klein, M. (1945). The Oedipus complex in the light of early anxieties, The Writings of Melanie Klein (pp. 370-419). London: Hogarth Press (1975).

Kohut, H. (1977). The Restoration of the Self. New York: International Universities Press.

Langner, T. F. (1962). A 22-Item screening score of psychiatric symptoms indicating impairment. Journal of Health and Human Behaviour, 3, 269-276.

Levinson, A., & Fonagy, P. (submitted). Attachment classification in prisoners and psychiatric patients. .

Luquet, P. (1981). Le changement dans la mentalisation. Revue Francais de Psychoanalyse, 45, 1023-1028.

Luquet, P. (1987). Penser-Parler: Un apport psychanalytique a la theorie du langage. In R.Christie, M.M.Christie-Luterbacher, & P.Luquet (Eds.), La Parole Trouble (pp. 161-300). Paris: Presses Universitaire de France.

Luquet, P. (1988). Langage, pensee et structure psychique. Revue Francais de Psychoanalyse, 52, 267-302.

Main, M. (1991). Metacognitive knowledge, metacognitive monitoring, and singular (coherent) vs. multiple (incoherent) model of attachment: Findings and directions for future research. In C. M. Parkes, J. Stevenson-Hinde, & P. Marris (Eds.), Attachment Across the Life Cycle (pp. 127-159). London: Tavistock/Routledge.

Main, M., & Goldwyn, R. (1994). Adult Attachment Rating and Classification System, Manual in Draft, Version 6.0. Unpublished manuscript: University of California at Berkeley.

Main, M., & Goldwyn, R. (1995). Adult attachment classification system. In M. Main (Ed.), Behavior and the Development of Representational Models of Attachment: Five Methods of Assessment : Cambridge University Press.

Marty, P. (1968). A major process of somatization: The progressive disorganization. International Journal of Psychoanalysis, 49, 246-249.

Marty, P. (1990). La Psychosomatique de l'Adulte. Paris: Presses Universitaire de France.

Marty, P. (1991). Mentalisation et Psychosomatique. Paris: Laboratoire Delagrangue.

McDougall, J. (1978). Plea for a Measure of Abnormality. New York: International Universities Press.

Mele, A. R. (1992). Springs of action: Understanding intentional behaviour. New York: Oxford University Press.

Morton, J., & Frith, U. (1995). Causal modeling: A structural approach to developmental psychology. In D. Cicchetti & D. J. Cohen (Eds.), Developmental psychopathology. Vol. 1: Theory and methods (pp. 357-390). New York: John Wiley.

Noam, G. G. (1990). Beyond Freud and Piaget: Biographical world - Interpersonal self. In T. E. Wren (Ed.), The moral domain (pp. 360-399). Cambridge, MA: MIT Press.

O'Brien, E. (1981). The self-report inventory: Construction and validation of a multidimensional measure of the self-concept. Sources of self-esteem, University of Massachusetts, Amherst.

Rogoff, B. (1990). Apprenticeship in thinking: Cognitive development in social contexts. New York: Oxford University Press.

Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect Grid: A single-item scale of pleasure and arousal. Journal of Personality and Social Psychology, 57, 493-502.

Segal, H. (1957). Notes on symbol formation,. International Journal of Psycho-Analysis, 38, 391-397.

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

Winnicott, D. W. (1962). Ego integration in child development. In D.W. Winnicott (Ed.), The maturational processes and the facilitating environment (pp. 56-63). London: Hogarth Press, 1965.

Winnicott, D. W. (1971). Playing and Reality. London: Tavistock.

Wollheim, R. (1995). The mind and its depths. Cambridge, Mass: Harvard University Press.

Reflective Functioning Manual: Outline of scoring procedures

General Rules:

1. Only explicitly reflective statements qualify for high ratings (referring to mental states is not enough -- there must be some demonstration of thinking *about* feelings or thoughts).
2. Learned, rote or clichéd statements do not qualify for high ratings.
3. Reference to personality or a relationship, in the absence of specific reference to mental states, does not qualify for high rating. Dynamic explanations and descriptions of personality, however accurate and perceptive they may seem, are not scorable as instances of high reflective functioning unless accompanied by specific references to mental states.
4. Giving benefit of the doubt must be carefully avoided. (Need to beware of halo effect here. There is a tendency to give credit for RF when a subject is likable and thoughtful, and this tendency to fill in the blanks in subjects' narratives must be avoided.) Similarly, it is vital not to rate somebody lower because they or their attitudes are unappealing. If their statements are reflective, they must be credited. To take a common example, a person who is egocentric or self-serving should be rated as reflective if they do make statements meeting criteria in the manual, *unless* the perceptions and explanations *in themselves* are self-serving.
5. Diagnoses should not be accepted as a short-hand for mental states.

Qualities which suggest moderate to high RF:

1. **Awareness of the nature of mental states**
(i.e., passages which demonstrate awareness of their **(1)** opacity, **(2)** susceptibility to disguise, and **(3)** potentially defensive nature; or which **(4)** demonstrate awareness of the limitations of insight into mental states, or which **(5)** make explicit reference to commonly expected reactions in specific situations.)
2. **Efforts to tease out mental states underlying behavior**
(Includes accurate attribution of mental states to others, recognition of diverse perspectives, taking into account how our own mental states affect behavior [ours and others'] and perceptions [our own and other's of us], etc.)
3. **Recognizing developmental aspects of mental states**
(Focus here is on how mental states change and evolve, and includes statements reflecting awareness of dyadic and family interactions. Note: awareness of intergenerational influences must contain explicit references to mental states and their influence on interpersonal behavior to count as +RF. Descriptions of interactions without understanding of the role of mental states is not scorable.)
4. **Showing awareness of mental states in relation to interviewer**
(Credit given for explicit efforts to clarify and help interviewer keep track of material, explicit and accurate references to the likely impact on interviewer of material a subject

has provided, statements demonstrating awareness that interviewer may not share subject's mental state in relation to one topic or another.)

Demand vs. Permit Questions:

Demand Questions -- must be rated

(Note: there is no penalty for non-reflective response if speaker has already responded to demand question in answer to previous question, or if a negative response seems plausible, e.g. no rejection described or few feelings about the loss of someone barely known to the subject):

1. Why did your parents behave as they did during your childhood?
2. Do you think your childhood experiences have an influence on who you are today?
3. (As a probe for influences of childhood experience) Any setbacks?
4. Did you ever feel rejected as a child?
5. (As a probe for losses) How did you feel at the time and how have your feelings changed over time? (Score separately for each loss.)
6. Have there been changes in your relationship with your parents since childhood?
7. Any demand-type question that an interviewer adds in a particular interview (i.e., "And why do you think they did that?")

Permit Questions -- all other questions: Note: Non-reflective responses to permit questions carry less weight than non-reflective responses to demand questions. Highly rated responses, however, should contribute to overall rating.

Guidelines for identifying and demarcating passages:

1. For responses to demand questions, give a single rating for the whole response, unless there are clearly different ideas contained within it. Rating should reflect the highest reflective level demonstrated in the response (though statements which are over-analytical or which show signs of negative reflection should be taken into account to reduce overall rating of passage).
2. For responses to permit questions (all other questions), score only if passage would receive a rating of '3' or higher (see below for criteria).
3. Instances of negative RF in response to permit questions can be taken into account when giving an aggregate score.

Guidelines for rating identified passages:

Note: All responses to demand questions must be scored, as well as relevant responses to permit questions.

-1 Negative RF

Response must:

- 1) be distinctly anti-reflective (i.e., hostile or actively evasive, usually because question is perceived as an assault or attack)
or
 bizarre (impossible to understand without making the assumption of irrationality on the part of the subject)
or
 inappropriate in the context of the interview (i.e., complete non-sequiturs over-familiarity, gross assumptions about the interviewer).

1 Absent but not repudiated RF

Response must:

- 1) be given in response to a demand question.
- 2) be passively rather than actively evasive.
- 3) be accompanied by little or no hostility.
- 4) contain no evidence of:
 - a) awareness of the nature of mental states;
 - b) explicit effort to tease out mental states underlying behavior;
 - c) recognition of the developmental aspects of mental states;
 - d) interaction indicative of the awareness of the interviewer's mental state
- 5) leave the interviewer no better off in terms of knowledge of the mental states of the subject, caregiver or other having read the passage than he/she was before reading it

Response may include:

- 1) concrete explanations of behavior in terms avoiding reference to mental states (i.e., explanations may be sociological, excessively general, or framed in terms of external, physical circumstances, etc.).
or
- 2) self-serving distortion (recollections which are highly egocentric, self-aggrandizing and/or contain extraordinarily arrogant claims to insight).

Note: The self-serving quality must be such that it leads the subject to make attributions that are clearly inaccurate and not simply biased or incomplete. Inaccurate efforts to tease out mental states underlying behavior are not sufficient to get a '1' rating unless they are *also* grossly self-serving.

3 Questionable or low RF

Response must:

- 1) contain some suggestion of mentalising efforts on the part of the subject which is nevertheless,
- 2) devoid of any element that makes reflective functioning explicit (i.e., it never reflects mixed emotions, conflict or uncertainty about beliefs and feelings of others).

Response may frequently:

- 1) make use of mental state language without making clear or explicit that the subject genuinely understands the implications of their statement.
- 2) appear somewhat clichéd, banal, superficial or 'canned.'
- 3) be excessively deep and detailed yet unconvincing and/or irrelevant to the task.

5 **Definite or ordinary RF**

Response must:

- 1) contain some feature which makes reflection explicit (i.e., explicit reference to the nature or properties of mental states, how mental states relate to behavior, or mental states in relation to the interviewer).
- 2) not be a cliché (though it does not need to reflect sophistication).

Response may:

- 1) show evidence of one of the six features (listed below) for assigning a rating of '7' in the context of a very simple observation of mental states which would otherwise rate only a '3.'

7 **Marked RF**

Response must:

- 1) contain some feature which makes reflection explicit (i.e., explicit reference to the nature or properties of mental states, how mental states relate to behavior, or mental states in relation to the interviewer).

and

- 2) meet at least one of the following. The passage:
 - ♦ is sophisticated (meeting at least 2 categories of qualities which suggest moderate to high RF).
 - ♦ is unusual or surprising, casting an original perspective (which is none-the-less readily understandable).
 - ♦ is complex or elaborate, described in unusual detail with indication that multiple mental states attributed to a person are considered in relation to one another.
 - ♦ places mental states within a causal sequence. Subject considers how the mental states arose, how they influenced behavior and what impact they have on subsequent perceptions, beliefs and desires.
 - ♦ provides evidence of an interactional perspective (outlining interactions of mental states between two people or within one person's mind).
 - ♦ contains an acknowledgment of a particularly painful situation, with appropriate thoughts and feelings.

9 **Full or exceptional RF**

Response must:

- 1) show the above features of '7 - marked RF' to an usually high degree (i.e., this response would be in the top 10% or less)

or

be given for a particularly charged and emotionally difficult subject in which maintaining even ordinary levels of reflective functioning could be considered exceptional.

- 2) have a strikingly personal character; it should enable the rater to feel confident that it is experienced as personally significant and meaningful.

Response may frequently:

- 1) demonstrate full awareness of important aspects of all protagonists within an interaction, such that the protagonists are placed in relation to one another in terms of

their feelings and beliefs and these are sufficiently complex and elaborate to convince the rater of their accuracy.

Rules for aggregating RF ratings into overall ratings

General Points:

- 1) Make a general judgment of the interview as a whole, rather than averaging scores on individual passages.
- 2) When confident that a particular transcript falls between two classes, assign the even number between those classes as an overall rating.

Chart of Overall Rating Criteria

Overall Rating	Common Types
-1 Negative RF <ul style="list-style-type: none"> ♦ subject systematically resists taking a reflective stance throughout interview. ♦ no passages rated '5' or above ♦ where some '1' or '3' passages exist, consider higher rating. 	A) Rejection of RF <ul style="list-style-type: none"> ♦ there are some general indices of neg. RF (i.e., lack of participation, hostility, evasiveness, marked incongruences). ♦ subject responds with hostile refusal to at least 3 demand questions. ♦ if subject gives only one or two hostile refusals, but meets general criteria above, rate '0'. B) Unintegrated, Bizarre or Inappropriate RF <ul style="list-style-type: none"> ♦ mental state attributions are confused and hard to understand. ♦ at least 3 examples of inexplicable, bizarre or inappropriate attributions (may occur in response to demand or permit questions) ♦ must be shocking rather than simply odd.
1 Lacking in RF <ul style="list-style-type: none"> ♦ reflective functioning is totally or almost totally absent. ♦ mental states may be mentioned, but there is no picture of the subject's or caregiver's beliefs and feelings underlying behavior. ♦ mentalisation is absent in the narrative and awareness of the nature of mental states, if present, not explicit. 	A) Disavowal <ul style="list-style-type: none"> ♦ barren accounts, lacking in mentalising detail. ♦ at least 3 examples of assertion of ignorance concerning mental states <i>or</i> comparable examples of evasion (physicalistic, behavioral or sociological accounts and global and generalized statements concerning psychological states) in response to demand questions. ♦ no instance of reflective function rated above '3' B) Distorted/self-serving <ul style="list-style-type: none"> ♦ interview does contain reflection, but reflection is flawed. ♦ reflective passages are egocentric, self-aggrandizing, and self-serving to the point where the accuracy of the representation of the mental state of the other may be reasonably called into question. ♦ at least 3 examples of such purposeful distortions in response to demand questions. ♦ no instance of reflective function rated above '5', or more than two above '3.'

Overall Rating**Common Types**

3 Questionable or Low RF

- ♦ Some evidence of consideration of mental states throughout the interview, but most references are not made explicit.
- ♦ Will contain some elements of a reflective stance.
- ♦ may contain more than 1 example of a rating of '5' or higher.
- ♦ must contain at least 3 examples of a '3' or '4' rating.

A) Naive/simplistic

- ♦ interview shows a partial understanding of intentions of others, but this understanding is likely to be banal, clichéd, and excessively general and superficial.
- ♦ normalization of experiences extends beyond what is culturally accepted.
- ♦ interview does not inter into complexities of mental states (conflicts, ambivalence, etc.
- ♦ naive/simplistic passages are the majority of low ratings.
- ♦ fewer than 3 ratings of '5' or above.

B) Over-analytical/hyperactive

- ♦ The interview may have greater depth than might be expected in the interview context.
- ♦ The interview is diffuse, however, and the insights are unintegrated.
- ♦ There are at least 3 instances in which the subject is over-analytical.
- ♦ If 1 or more of these includes statements that are bizarre, distorting/self-serving, consider '1' or '2' rating.

C) Miscellaneous low RF

- ♦ transcript is neither particularly naive nor overly analytic.
- ♦ this rating may be a compromise between ratings for transcripts which show marked disavowal mingled with definite evidence of reflective functioning (or other such incongruities).

5 Ordinary RF

- ♦ There are a number of instances of reflective functioning (and these may be prompted, rather than spontaneous).
- ♦ Speaker has a model of the mind (own and attachment figures) which may be simple but is relatively coherent, personal, and well-integrated.

A) Ordinary Understanding

- ♦ Subject shows an ordinary capacity to make sense of their experience in terms of thoughts and feelings.
- ♦ Subject has a consistent model for thoughts and feelings of self and other which requires little or no inference from the rater.
- ♦ This model is limited, and does not include understanding of conflict or ambivalence.
- ♦ There are at least 3 passages rated '5.'
- ♦ No breakthroughs of rejection, bizarre explanations, pervasive disavowal, etc.

Overall Rating**Common Types**

5 Ordinary RF (Cont.)

- ♦ Must have at least 1 or 2 clear '5' passages. Most interviews with this rating will have responses in the '3' to '7' range.
 - ♦ If any '-1' or '1' ratings, these are balanced by passages immediately following or elsewhere which indicate reflection.
- B) Inconsistent Understanding**
- ♦ One or two passages warrant a '7' rating or even higher, but understanding cannot be maintained in relation to one or more problem areas (i.e., a conflictual relationship to one parent).
 - ♦ Even problematic parts of interview do not fall below a '1' or '2' rating.

Overall Ratings for high RF (no sub-types)

7 Marked RF

- ♦ Numerous instances of full reflective functioning suggesting a stable psychological model of the mind (own and caregivers') and reactions to mental states.
- ♦ Usually, passages where subject has arrived at an original reintegration of states of mind (own and/or others).
- ♦ Much detail about thoughts and feelings
- ♦ Implications of mental states explicitly spelled out.
- ♦ Usually able to maintain a developmental (interactional) perspective.
- ♦ In interview as a whole, subject is applying reflective stance fairly consistently to at least one context, or less consistently to a number of contexts.
- ♦ At least 3 instances, anywhere in interview, which rate '7' or higher.
- ♦ No passages rated '1' or lower, unless no relevant experiences to rate.
- ♦ No more than 3 passages where rating is less than '5' in response to demand questions, where there are relevant life experiences.

9 Exceptional RF

- ♦ Transcript shows exceptional sophistication, is commonly surprising, quite complex or elaborate and consistently manifests reasoning in a causal way using mental states.
- ♦ Shows consistent reflective stance across all contexts.
- ♦ Has 3 or more instances, anywhere in interview, with a '9' rating (i.e., the response integrates several aspects of reflective functioning into a unified, fresh perspective).
- ♦ Few passages rated '3' and most would be rated '5' or '7.'
- ♦ If above criteria are not met but rater "feels" the transcript to be exceptional, a rating of '8' should be considered. (For an '8', should be no more than a couple of passages rated '3' and more than one '8' or '9'.)