Reflective Functioning on the Parent Development Interview: Validity and reliability in relation to socio-demographic factors

Michelle Sleed[1,2], Arietta Slade[3] & Peter Fonagy[2]

[1]Child Attachment & Psychological Therapies Research Unit (ChAPTRe), Anna Freud National Centre for Children and Families, London, UK

[2]University College London, London, UK

[3]Yale Child Study Center, New Haven, CT, USA

Corresponding author:

Michelle Sleed, Child Attachment & Psychological Therapies Research Unit, Anna Freud National Centre for Children & Families, 12 Maresfield Gardens, London, NW3 5SD, United Kingdom

Michelle.Sleed@annafreud.org

Word count: 7991

REFLECTIVE FUNCTIONING ON THE PARENT DEVELOPMENT INTERVIEW: VALIDITY AND RELIABILITY IN RELATION TO SOCIO-DEMOGRAPHIC FACTORS

Abstract

Background: The Reflective Functioning coding of the Parent Development Interview (PDI-RF) is a widely used method for assessing a caregivers' capacity for mentalization. However, little is known about its psychometric properties.

Aim: This study examined the distributions and discriminant and criterion validity of the PDI-RF scale in relation to a number of demographic and socioeconomic factors.

Method: Mothers of infants and toddlers (N = 323) from low, medium and high-risk samples were interviewed with the PDI and transcripts were coded for RF. Demographic and socio-economic data were recorded.

Results: The PDI RF scale showed high inter-rater reliability, internal consistency, and criterion validity. Modest associations with some sociodemographic variables and PDI RF were found, but together these only accounted for a small amount of variance in the measure, suggesting adequate discriminant validity. Overall, the scale had good psychometric properties, although some caveats for its use were identified.

Keywords: Parent Development Interview, PDI, Reflective Functioning, Psychometric Properties, parental mentalizing

Introduction

Mentalization refers to the process of understanding how mental states - feelings, beliefs, desires, intentions, and knowledge – affect individuals' perception of their own and others' behavior (Fonagy, Gergely, Jurist, & Target, 2002; Fonagy, Target, Steele, & Steele, 1998; Sharp & Fonagy, 2008). Reflective Functioning (RF; Fonagy, Steele, Steele, Moran, & Higgitt, 1991), one of the best operationalizations we have of this process, measures the ability to read and interpret behaviors in light of the mental processes that underlie them and, at the most general level, reflects an individual's ability to understand others as intentional agents and see themselves in the same manner.

The development of this capacity is a normal yet crucial part of human socio-emotional development. Making meaning of one's own subjective experiences in terms of mental processes allows for a deep and broad knowledge of one's self in relation to affective states and forms the vital structures necessary for affect regulation (Slade, 2005). The capacity to do so emerges within the context of the parent-child relationship (Fonagy, et al., 1995). It is the parent who first makes meaning of the child's experience, setting the stage for how he will come to know himself. In infancy, the parent is the first 'decoder' of the child's inner world, finding sense and coherence in his behaviour. She conveys that sense to him through her behaviour, such that his feelings and thoughts are brought to life in their interactions. Her understanding of the infant will guide her behaviour and allow her to respond sensitively; for example, she sees his reticence in a new situation as slight fear that requires her comfort.

Ideally, parents begin to make meaning of their babies in earliest infancy. These efforts – to consider the mental states of the infant, and to make sense of their behaviour give shape to the baby's inner world; the child comes to know himself as he is seen and known by

his parents.   The more the child is seen by the parent as meaningful and coherent, the more

he will come to experience himself as meaningful and coherent (Fonagy & Target, 1997;

Slade, 2005). The parent's ability to consider the mental states of the infant is linked with the

mother's sensitivity in interaction, the security of the child's attachment, and the emerging

capacity to regulate his/her affective states (Grienenberger, Kelly, & Slade, 2005; Slade,

Grienenberger, Bernbach, Levy, & Locker, 2005). As parental mentalization is thought

essential to a range of developmental and socio-emotional processes, impingements on its

development, particularly as the result of adverse childhood experiences, will generate

maladaptive (psychopathological) outcomes (Fonagy, 2003). Evidence consistent with this

developmental model has been accumulating, for example, in investigations of the impact of

severe childhood trauma on RF (Berthelot et al., 2015; Ensink, Bégin, Normandin, & Fonagy,

2015; Ensink, Berthelot, Bernazzani, Normandin, & Fonagy, 2014; Ensink, Fonagy,

Berthelot, Normandin, & Bernazzani, 2015; Ensink, Leroux, Normandin, Biberdzic, &

Fonagy, 2015; Newbury-Helps, Feigenbaum, & Fonagy, 2016; Sharp et al., 2016).

Mentalization research has been strongly facilitated by an objective narrative-based

measure of RF. Fonagy and colleagues developed a system for quantifying a person's

capacity for RF (Fonagy, et al., 1991; 1998) based on their responses to questions on the

Adult Attachment Interview (AAI; George, Kaplan, & Main, 1984). This is a semi-structured

interview that asks adults to talk about their childhood relationships with their

parents/caregivers with a remarkable capacity to predict the quality of child-parent

attachment (Bakermans-Kranenburg & van IJzendoorn, 2009; Verhage et al., 2016). An

addendum to the original RF coding system (Slade, Bernbach, Grienenberger, Levy, &

Locker, 2004) was developed for use with narratives on the Parent Development Interview

(PDI; Slade, Aber, Bresgi, Berger, & Kaplan, 2004), a semi-structured interview that taps

into parental representations of the child, themselves as parents, and of the parent-child

relationship. It elicits similar narratives to the AAI in that it explores emotionally charged attachment relationships. However, while the AAI taps into adults' relatively stable representations of past relationships, the PDI asks about current relationships that are dynamic and developing (Slade, 2005).

The validity and reliability of the PDI RF coding system

While some evidence for the validity of the RF rating on the AAI has been established (Arnott & Meins, 2007; Berthelot et al., 2015; Bouchard, Target, Lecours, Fonagy, & Tremblay, 2008; Chiesa & Fonagy, 2014; Ensink, Bégin, et al., 2015; Ensink, et al., 2014; Ensink et al., 2015; Fischer-Kern et al., 2010; Gullestad, Johansen, Hoglend, Karterud, & Wilberg, 2013; Ha, Sharp, Ensink, Fonagy, & Cirino, 2013; Katznelson, 2014; Pedersen, Lunn, Katznelson, & Poulsen, 2012; Rutherford, Goldberg, Luyten, Bridgett, & Mayes, 2013; Steele & Steele, 2008; Vermote et al., 2010), relatively little is as yet known about the validity of RF as measured on the PDI, especially in relation to socioeconomic and demographic factors. The research that has been carried out with the PDI RF coding system has either looked at RF in relation to maternal and infant attachment status (Slade, Grienenberger, Bernbach, Levy, & Locker, 2005; Stacks et al., 2014), maternal behavior (Grienenberger, et al., 2005; Schechter et al., 2008; Stacks et al., 2014), improvements in maternal attributions (Schechter et al., 2006), maternal psychopathology (Schechter, 2003; Schechter et al., 2005), or as a treatment outcome (Fonagy, Sleed, & Baradon, 2016; Pajulo et al., 2012; Sadler et al., 2013; Sleed, Baradon, & Fonagy, 2013; Suchman et al., 2010; Suchman, Decoste, McMahon, Rounsaville, & Mayes, 2011; Suchman, Legow, Decoste, & Castiglioni, 2008). These studies have, to a certain extent, demonstrated the concurrent and predictive validity of the measure of RF on the PDI and its sensitivity to treatment change. However, since no research into the discriminant validity has as yet been reported, little is

known about how the measure relates to the characteristics of the parents being interviewed.

It is possible that a number of demographic and socioeconomic characteristics may be confounding variables in the measurement of parental RF. For example, imagining the intentions and feelings of a two-month old infant would likely be more difficult than imagining the subjective experience of an 11-month old baby or three-year old child. And yet, mentalization theory has emphasized the importance of the mother's capacity to mentalize in the very early months of life (Fonagy, Gergely, & Target, 2007; Sharp & Fonagy, 2008). It is therefore important to determine whether the coding system is valid for use with parents of very young infants. Similarly, it is important to know whether the system is valid for mothers across a range of ages.

A further question is whether measures of RF based on verbal narratives elicited by semi-structured interviews such as the AAI and PDI are actually measuring the person's ability to describe him/herself coherently and logically, rather than to understand mental states of self and other. In other words, is the RF coding system measuring some aspect of education or intelligence rather than the ability to mentalize? Recent research has shown parental RF on the PDI to be related to both IQ and executive functioning in a sample of substance-dependent mothers (Håkansson, Söderström et al. 2018). Further research is needed to see if these same findings are replicated in other high-risk and low-risk samples of parents.

Most of the research that has been carried out with the PDI RF coding system has focused on specific samples of high risk families who are likely experiencing relatively high levels of social deprivation. The extent to which various socioeconomic stressors and family demographic characteristics relate to RF ratings in a wider and more diverse sample of parents is not yet known.

The aim of the current study was to examine the distributions and psychometric properties of PDI RF ratings in a relatively large sample of normative and high-risk parent-child dyads. The research questions were:

1) What are the distributions of PDI RF ratings in a large and diverse sample of mothers of infants and toddlers?

2) What is the internal consistency and factor structure of the PDI RF ratings for individual questions and overall scores?

3) What is the discriminant validity of the PDI RF coding system in relation to a number of demographic and socioeconomic factors?

4) What is the criterion validity of the PDI RF coding system in relation to different parenting risk groups?

The data were drawn from three samples: a non-referred community group ("normative sample"), a group of clinically referred dyads ("clinical sample"), and a group of mothers and babies in prison. On the basis of prior research we expected that RF would be lower in the clinical than the normative group and would be lower still in the prison group. We predicted that while demographic variables are likely to be associated with RF, differences in RF between the clinical and normative group and the prison and normative group would remain even after demographic variables were controlled for.

Method

Participants

The sample for this study (N = 323) was drawn from three separate studies: a clinical study of mothers with mental health problems who had young babies (clinical group; n = 118), a non-

clinical community sample of mothers with young babies (normative group, n = 56), and a study of mothers and babies staying on Mother Baby Units (MBUs) in prisons (prison group, n = 149). The clinical and prison samples were drawn from randomized controlled trials (RCT) of intervention effectiveness. The normative group was recruited as a non-clinical comparison group for this study. Only baseline pre-intervention data was used for the purposes of this study.

Sample 1: Clinical group

Mothers in this group were participating in an RCT comparing the outcomes of Parent-Infant Psychotherapy with Treatment as Usual for mothers with mental health problems and their infants (Fonagy, Sleed, & Baradon, 2016). These participants were recruited from inner city areas with high levels of social exclusion and deprivation. Mothers were identified by their psychiatrist, health visitor, midwife, GP, or other professional as requiring additional emotional support. Families were eligible for inclusion for the therapeutic trial if, a) mothers met probable psychiatric caseness criteria based on the General Health Questionnaire (GHQ; Goldberg, 1978), b) their infant was less than 12 months of age, and c) mothers met at least one of the following further indicators of social exclusion: low income household; long-term unemployment (longer than 2 years); living in temporary or overcrowded accommodation; single or unpartnered; chronic physical illness or disability; early childhood history of foster or institutional care; social isolation associated with recent relocation; less than 20 years of age; or previous diagnosis of non-psychotic psychiatric illness. Mothers were excluded from the sample if they were non-English speaking, had a current diagnosis of psychosis, had substance abuse disorders/chronic drug dependence or an IQ below 70.

Sample 2: Non-clinically referred (normative) group

Participants were recruited from mother and baby groups and children's centers in inner-city

areas where there were comparable levels of social exclusion and deprivation to that of the clinical group. In contrast with the clinical sample, this sample of mothers had not been identified by a professional as having mental health problems, and they did not meet the GHQ cutoff. This study had the same exclusion criteria that were applied to the clinical sample above.

Researchers gave information to mothers whose infants were under 12 months of age attending the children's centers, and participation was on a voluntary basis. Families who were potentially interested in taking part gave their contact details to the researcher. These mothers were then contacted by telephone, and if they agreed, an appointment was made. A small financial incentive was offered as part of the invitation to participate.

Sample 3: Prison group

Participants in this group were mothers and babies who were participating in a cluster RCT evaluating the outcomes of an intervention in MBUs in UK prisons (Sleed, et al., 2013). Mother-baby dyads currently on MBUs in the participating prisons were eligible to take part in the project and were invited to participate. Dyads were excluded if they were not fluent in English or were due to leave the unit before the follow-up interviews. Before the evaluation took place on any unit, the researcher (and, in intervention sites, the course facilitator) met with unit staff involved in recruitment of participants to the project. Mothers in the intervention group were recruited from three MBUs and mothers in the control groups were recruited from the 4 MBUs comparable in size and demography that were not running the intervention. The research psychologist then visited the MBU one week before the start of the course to give an introductory talk about the research to those mothers identified as being eligible and willing to participate and arranged to see each mother separately. It was made clear to each participant that they were free to leave the research at any time and that doing so

would have no implications for their involvement on the course or their sentence. All data included in this analysis were collected at baseline. The parents and children in the control group were matched, as far as possible, with the intervention group for mother's age, child's age and gender.

Description of the total sample

The demographic characteristics of the three subsamples are summarized in Table 1. The mothers in the study were aged between 18 and 50 years old, with most of them in their late 20's or early 30's (mean = 30). The target children were from newborn to 2 years of age, averaging about 5 months old. There were almost equal numbers of male and female children, and more than half of them were the mothers' only children. The mothers in the full sample were representative of a very broad range of educational and ethnic backgrounds. The three groups differed significantly in terms maternal age, $F(2) = 37$, $p < .001$, child age, $F(2) = 12.9$, $p < .001$, child gender, $X^2 (2, N = 323) = 8.2$, $p = .017$, ethnicity, $X^2 (8, N = 323) = 33.7$, $p < .001$, and maternal education, $X^2 (6, N = 323) = 139.0$, $p < .001$. Post hoc tests revealed that mothers in the prison sample were younger, less educated, had more female babies and more families of Black ethnicity. The children in the normative group were significantly older than those in the two other samples, and mothers in the normative group had a significantly higher nonverbal IQ than mothers in the clinical group ($t_{(160)} = -2.99$, $p = .003$. There were no differences between the three groups in terms of the number of other children the mother had, $X^2(10, N = 323) = 16.0$, $p = .100$, NS.

(Insert Table 1 here)

Social exclusion criteria

The clinical and normative study designs stipulated that each family in the study met at least one of a list of nine social exclusion indicators. These are listed in Table 2.

(Insert Table 2 here)

The families in the clinical group were likely to meet more of the social exclusion criteria than those in the normative group. More specifically, the mothers in the clinical group were more likely to be on income support (X^2 (1, N = 174) =16.9, p < .001), unemployed for more than two years (X^2(1, N = 174) = 10.5, p = .001), living in temporary or overcrowded accommodation (X^2 (1, N = 174) = 4.6, p = .032), single (X^2 (1, N = 174) = 17.4, p < .001), or to have had a previous psychiatric diagnosis (X^2 (1, N = 174) =35.7, p < .001). The two groups did not differ in terms of maternal physical illness or disability, maternal foster or institutional care, social isolation, or whether or not mother was a teenager.

Ethical considerations

Ethical approval for the three studies from which the sample was drawn was granted by the appropriate boards for each study. All participants gave informed consent and data were stored and handled according to the UK Data Protection Act 1998.

Measures

All measures were administered by researchers in the settings from which participants were recruited.

PDI (Slade, Aber, Bresgi, Berger, & Kaplan, 2004)

The Parent Development Interview (PDI) is a semi-structured clinical interview that taps into parents' representations of themselves as parents, of their child, and of the relationship between them. The interviewee is asked to focus on a current and specific relationship with one child. The parent is asked to give a number of adjectives describing her child, herself as a mother, and their relationship, and to give examples for these. She is also asked about feelings of joy, pain, guilt and anger as a parent, and to consider the child's feelings in a number of ways. Another set of questions relate to the parent's thoughts and feelings about her own and the child's feelings during separations. There are also several questions about how the parent feels she has been affected as a parent by her experiences with her own caregivers.

The interviews were transcribed verbatim and coded with the RF coding system (Fonagy et al., 1998) using an addendum that was specifically adapted for this interview (Slade, Bernbach, et al., 2004). The PDI-RF coding addendum has the same 11-point RF rating scale and descriptions of RF types as the original coding, but these are described and exemplified specifically for narratives about current parent-child relationships.

Two types of scores are assigned. First, a subset of "demand" questions are coded for level of parental RF. These questions were chosen because they require the parent to mentalize in order to respond, i.e. to describe what they, their child, or both of them may have thought or felt. The remaining questions, while not specifically scored, also provide useful insights into parental representations and reflective capacities. On the basis of individual demand question scores, as well as a reading of the entire interview, the coder then assigns an overall RF score. The RF scale has a potential range of -1 (negative or bizarre RF) to 9 (marked RF). Scores between 4 and 6 are considered moderate, scores of 3 or below are

considered low, and scores of 7 or above are considered high (Fonagy, et al., 1998). This score range applies to the demand questions as well as the overall score.

Test of Non-verbal Intelligence (TONI-III; Brown, Sherbenou, & Johnsen, 1997)

The clinical and normative groups were assessed with the TONI-III. This test is a language-free measure of cognitive ability that is robust and highly predictive of general intellectual functioning, unaffected by cultural differences, and has good reliability and validity (Atlas, 2001; Banks & Franzen, 2010; Brown, et al., 1997). Mothers were shown a series of patterns and asked to identify the missing image in the sequence from a multiple choice set of answers. The raw total scores were converted to standardized estimates of nonverbal intelligence based on large population norms. These scores are standardized with a mean of 100 and a standard deviation of 15.

Results

Distribution of PDI RF ratings

The details of the RF scores for each demand question and the overall interview are presented in Table 3. The mean RF scores for all three sub-samples were, for most of the demand questions, moderate to low. In many cases the full range of potential scores was not used; no parent scored at the lowest (-1) or highest (9) level for some questions.

(Insert Table 3 here)

Tests of normality showed that the RF scores for demand questions in each of the subsamples were non-normally distributed (Kolmogorov-Smirnov test, p< .005). The only exceptions were the "guilty" and "angry" questions, for which the distribution of scores in the normative

sample was normal. The histograms and Normal Q-Q plots for each demand question in each

of the subsamples were examined, and the prison sample had more skewed distributions than

the two community samples. This was particularly the case with negatively framed questions

("not clicked", "angry" and "rejected"), as the mothers in the prison sample were highly

likely to respond with a disavowal of such emotional states, scoring at the low end of the

scale. This pattern of distributions may lead to a floor effect in the measurement of these

domains in this population.

Inter-rater reliability

The PDIs in the full sample were coded by seven coders. Each coder was trained by a

certified expert in coding RF on the PDI (the first author). As part of the accreditation

process, each rater coded a reliability set of ten interviews and each coder attained a good

reliability score with the author of the interview and coding system (ICC > .75).

For this study, a subset of 17 PDI interviews was double coded by the RF raters who

contributed to this data set. The intraclass correlation co-efficient for the overall RF score

was high (ICC = .87).

Internal consistency

The internal consistency of the 15 demand individual questions was examined. The

Cronbach's alpha was high, (α = .90), and remained relatively stable if any of the items were

removed. If the overall RF score was included in the scale, the Cronbach's alpha increased

further (α = .91). This pattern of results was similar for each of the three subsamples when

examined separately.

Factor structure of PDI RF ratings

The correlations between the RF scores for each demand question and the overall RF scores are presented in Table 4. All correlations were highly significant (p < .001). The question that was mostly strongly associated with the overall score was the "parent" question. The least strongly associated demand question was the "losing" question. This potentially triggering question was omitted in the prison sample, because many of the mothers faced separation from their babies if the babies aged out of the program (at 18-24 months) before mothers completed their sentences. A principle components analysis of the RF scores for the individual demand questions was carried out to examine the factor structure of the items. A single factor solution best fit the data, confirming that all scores contribute relatively equally to the overall coding system.

(Insert Table 4 here)

Discriminant validity

In order to test discriminant validity, we examined the correlations between the overall RF scores and a number of socio-demographic variables (child gender, child age, maternal age, first time parenthood, maternal education, IQ, and social exclusion variables) in order to determine whether any of these contributed to variations in RF across the three maternal risk groups (see Table 5).

    Infant variables. Child gender was not significantly correlated with RF. Child age was not correlated with RF scores in the clinical or normative samples, but in the prison sample, mothers of older children tended to have slightly higher RF scores. Note that the variability of child ages in the other two sub-samples was quite restricted, while child age the prison sample ranged from 2 weeks to two years old. To examine the threshold at which

15

maternal RF is potentially confounded by infant age, a number of dummy variables of infant age cut-off points were created. Taking into account the full sample, the overall RF rating for parents of infants younger than 2 months was significantly lower than those whose children were more than 2 months old, t (318) = -2.32, p = .021, d = -.29. This difference was not significant if a cut-off of any other age between 3 and 12 months was used.

Maternal variables. The age of the mothers was not related to RF scores within the subsamples, but when the whole sample was examined, older mothers had significantly higher RF ratings. As the mothers in the prison sample were younger than those in the other groups, a partial correlation was carried out to test if the relationship between maternal age and RF was still significant after controlling for group (prison vs. others). Although the strength of relationship was lower, the association was still significant (r = .11, p = .050). First time mothers in the clinical sample were significantly more likely to be reflective than those with other children. While this association was not significant for the other two subsamples, the association for the combined group was significant.

Mothers who were educated beyond high school level had significantly higher RF scores in the prison sample; this association was also significant for the overall sample. To examine whether this relationship could be explained by IQ, further analyses of the normative and clinical samples were carried out (data on IQ was not available for the prison group). For these two subsamples together, RF and maternal education beyond high school level were significantly correlated (r = .20, p = .008). When a partial correlation controlling for IQ was carried out, the strength of association was no longer significant at the $p < .05$ level (r = .14, p = .089). Thus, there is some modest evidence that education level is related to RF ratings, but this is partially explained by IQ.

Data on maternal nonverbal IQ and social exclusion indicators were only collected in the clinical and normative samples and only data from these samples is presented below. The

correlation between maternal RF and nonverbal IQ was significant in the clinical sample and the pooled sample (see Table 5). This may indicate that nonverbal IQ levels at certain levels were more likely to confound the measure of RF (since the mean nonverbal IQ scores were significantly lower for the clinical group than the normative group). Dummy variables of IQ levels above and below a range of cutoff scores were used to determine if the association between RF and IQ pertained only to particular levels of IQ. The TONI-III is standardized to have a mean of 100 and a standard deviation of 15. The different cutoff points examined were: one SD below the mean of the test (85), the test mean (100), the actual mean of this sample (107), and one SD above the mean of the test (115). The level of overall RF was significantly lower for those mothers whose nonverbal IQ levels were below 85 than those with higher nonverbal IQ levels, $t_{(156)} = 2.25$, $p = .026$, $d = .74$. Equally, mothers with a nonverbal IQ below the test mean of and below 100 had significantly lower levels of RF than those with higher IQ levels ($t_{(156)} = 3.20$, $p = .002$, $d = .62$). However, overall RF levels were no longer significantly different when comparing mothers with nonverbal IQ levels above or below the sample mean (107), or when higher cutoff IQ scores were used. These findings suggest that lower than average levels of nonverbal IQ were likely to confound the measure of RF.

In both normative and clinical groups, long-term maternal unemployment was significantly correlated with RF. The only other indicator of social exclusion that was related to maternal RF was if the family was eligible for income support, and this was only the case in the clinical group. When a partial correlation was carried out between RF and income support eligibility, controlling for long-term unemployment, this correlation was no longer significant ($r = -.15$, $p = .083$). In other words, long-term unemployment is the social exclusion factor most strongly related to RF and the one that mostly explains the apparent link between income support eligibility and RF.

17

(Insert Table 5 here)

Stepwise linear regression analyses were carried out to investigate which socio-demographic factors were predictive of maternal RF (Table 6).

(Insert Table 6 here)

The variables that showed significant correlations with RF and were available for the full sample were selected as independent variables. These were: child age, mother age, mother educated post high-school level (dummy variable), and the number of children the mother had. The four predictor variables were entered into a stepwise linear regression model. In the final model, all predictors but child age were significant. This model accounted for just over 10% of the variance in RF ($R^2 = .13$, $F(3, 311) = 17.04$, $p < .001$).

There were some modest correlations between the predictor variables in the final model (as shown in Table 7) suggesting that multicollinearity may be an issue. However, the tolerance statistics for all predictors were well above .01 (range .84 to .92) and the Variance Inflation Factors were all > 1 and <10 (range 1.08 to 1.19), suggesting little threat from multicollinearity in the final model (Mansfield and Helms, 1982).

(Insert Table 7 here)

 Criterion validity

In order to examine whether or not the RF scale differentiated between mothers in the three different risk groups, a one-way analysis of variance was carried out. As hypothesized, mothers in the three groups had significantly different overall RF scores ($F(2,319) = 18.46$, p

< .001). This remained the case even after controlling for the possible confounding variables identified in the above analyses (child age, mother age, first time motherhood, maternal education). A post hoc Tukey HSD test revealed that the overall RF scores were different for all three groups. Mothers in the prison sample had RF scores significantly lower than those in the clinical ($p < .001$, $d = 0.24$) and normative ($p < .001$, $d = -0.88$) groups, and mothers in the clinical group had RF scores lower than those in the normative group ($p =. 037$, $d = -0.41$).

Discussion

This study has provided the first analysis of a large number of PDIs coded for maternal RF, and thus contributes in a number of ways to our understanding of the psychometric properties of the PDI RF coding system. On the whole, the coding system demonstrated high levels of inter-rater reliability, internal consistency, and solid evidence of discriminant validity. The PDI RF system also distinguished between groups with different levels of risk, thus establishing the measure's criterion validity. In particular, mothers in a normative community sample had higher RF than mothers in a clinical sample, who in turn had higher RF than mothers in a prison sample. Taken together, these results suggest that the PDI and the accompanying PDI RF coding system have strong psychometric properties, which can – with some caveats – be used in research with diverse typical and clinical populations.

Distribution of the PDI RF coding system

The distribution of RF scores tended to be non-normally distributed for the full sample in this study. For many of the demand questions as well as the overall RF scores, the full range of scores was not used. This was particularly the case with the most high-risk cohort, the

19

mothers in prison. Thus, despite having a relatively heterogeneous pooled sample in this study, no parents were rated overall as anti-mentalizing or having bizarre of hostile representations (scoring -1), or as being exceptionally reflective (scoring 9). Other studies of RF in high risk populations (Pajulo et al., 2012; Smaling, et al., 2015; Suchman, et al., 2011) have likewise reported non-normal distributions of RF, whereas RF scores in low-risk samples tend to be more normally distributed (Slade et al., 2005; Stacks, et al., 2014). None of the samples studied here could be described as low-risk, as even the non-clinical population in this study was drawn from inner city sites with relatively high levels of social deprivation. At the same time, none of the samples was at the highest risk: even the mothers in the most high-risk prison sample had been granted permission by a multidisciplinary board to remain on the MBUs with their children following a rigorous risk assessment. Had we sampled mothers who were deemed to be of great risk to their babies, such as those whose children were on the Child Protection Register or families going through child care proceedings, we might have tapped into the extreme lower end of the scale. In any event, both our findings as well as those of other researchers suggest that when working with high risk samples assumptions of normality should be evaluated carefully.

Inter-rater reliability and internal consistency of the RF coding system

Two assessments of reliability were carried out in this study, measures of inter-rater reliability and internal consistency. The inter-rater reliability of the coding system was high, both on the training set and as a measure of inter-rater reliability for the current study.

The internal consistency of RF ratings for all of the demand questions was also very high, as were the correlations between scores for each of the demand questions with the overall RF score. This means that one can confidently use the overall score as a single indicator of the parent's mentalizing capacity, as is suggested by the author of the coding

20

system (Slade, Bernbach, et al., 2004). Interestingly, the one demand question in the PDI that is not specific to the parent-child relationship, the question about the mother's childhood experiences with her caregivers, is the one that correlated most strongly with the overall RF rating. It may be that this question is particularly good at evoking a sense of her capacity to mentalize about attachment relationships in general.

Discriminant validity of RF on the PDI

As we discuss below, initial correlation and regression analyses indicated that PDI-RF was related to a number of socio-demographic variables, including child and maternal age, the number of children a mother has, her level of education and nonverbal intelligence, and her history of unemployment. However, stepwise linear regressions including potential confounding factors indicated that these account for a small amount of the variance in predicting RF, thus establishing the discriminant validity of the PDI RF coding system.

Infant variables.

In the sample at large, but especially in the prison sample, mothers of older infants tended to have higher levels of RF. Subsequent analyses revealed that the RF scores were significantly lower for mothers of infants two months or younger than mothers of older infants. Thus, mothers of newborns and very young infants find it difficult to think of their child's internal experience in particularly differentiated ways. Given that the infant is very new to them, this makes intuitive sense. After all, they have just met. This is not to say that parents are not engaged in mentalizing processes from birth onward; it is in fact in the early months of development that this mentalization process is most important for the infant's developing sense of self and capacity for self-regulation (Fonagy et al., 2002). Even very young infants have affective experiences that need to be thought about by the parent,

represented in her mind, and re-presented back, in behaviour and language. But the limitations of a parent's experience with the baby in the first two months likely makes accurate and nuanced reflections challenging. It is also possible that the questions in the PDI do not adequately capture the experience of parenting a newborn baby. These findings suggest that the coding system may not be an appropriate measure to be used with parents of infants younger than two months. It does, however, appear to be robust for parents of infants above two months of age; it is at this time that babies begin smiling and engaging reciprocally, making them much easier to know and read.

Research has shown that infants start to attribute goals and intentionality from about 7 months onwards (Csibra, 2008; Kovács, Téglás, & Endress, 2010). This level of intentionality and representation makes the interplay between the parent's and infant's minds more complex and therefore may make reflective functioning more explicit in narratives from parents of older infants. However, the current study demonstrated that parents – beginning when their infants are as young as two months old - are attempting to work out what their infant may be thinking or feeling, and to consider their own thoughts and feelings within the relationship.

The psychoanalytic literature also supports this finding. According to early parent-infant theorists, in the weeks just before and after the birth of a baby, mothers enter into a state of "primary maternal preoccupation" (Winnicott, 1956), or the "motherhood constellation" (Stern, 1995), a particular state of mind characterized by the mother's preoccupation with caring for her baby, who she at first represents as somewhat undifferentiated from her own self. Using her own experiences as a baby, the mother is thought to become very much identified with her infant and acutely aware of what he or she may be feeling. During this very early phase of the parent-infant relationship, it is possible that the mother's capacity to differentiate her own mental states from those of her baby is

22

more difficult. What may be seen as a lack of mentalizing about the infant as an individual person may actually be the mother's adaptive identification with her baby which is a necessary precursor to her later capacity to make sense of her baby's internal world in a more differentiated way. It is also likely that the potent need for a mother to understand and regulate her newborn's and her own physical states - hunger, tiredness and pain - takes precedence in her representation of the relationship in the early postnatal period. These concrete themes may dominate her narrative, seemingly at the expense of more internal mental state references (although she may still have an emerging model of the infant's mind in her mind). This may continue until the baby has established more regulated eating and sleeping patterns and is better able to communicate these physical needs in a differentiated manner, enabling the mother to shift her focus to the representation of mental states.

Maternal variables.

Older mothers tended to have higher levels of RF, even when the relative youth of the prison sample mothers was controlled for. This is very much in line with the fact that the capacity for RF develops over time and across adolescence and early adulthood (Burnett & Blakemore, 2009). In addition, first time mothers tended to have higher levels of RF than those with other children, particularly in the clinical sample. This is contrary to what might be expected - that the experience of parenting leads to a better capacity to understand and consider a baby's thoughts and feelings. However, this study showed that having multiple children was associated with a poorer capacity to mentalize. There may be a number of explanations for this. First, the inexperienced first time mother may need to work harder to make sense of her young baby's experience; this active process of mentalizing would thus be more evident in the interview than it might be with more experienced mothers of infants. Second, a mother with several children, particularly when they are close to each other in age, is likely to be experiencing high levels of stress. As many have noted, stress and distress

inhibit mentalizing abilities (Fonagy et al., 2002), and thus, potentially, the RF of mothers who are overwhelmed with caregiving. Related to this is the fact that when children are closely spaced, really being able to think about each one individually can be challenging, and the differentiation of components of emotional experience is a key component of successful mentalizing.

Lower levels of maternal education and nonverbal IQ were both associated with lower levels of RF, although most of the variance was explained by IQ. The association between maternal nonverbal IQ and RF is unsurprising, and – in our view – not a particularly strong indication that IQ confounds the measurement of RF on the PDI. Language based assessments of mentalization, a metacognitive capacity, likely require some capacity for abstract reasoning and problem solving, which is what the nonverbal IQ assessment is measuring (Brown, et al., 1997). In addition, PDI RF requires a basic level of cognitive capacity to both understand questions and provide meaningful responses. Previous research of RF on the Adult Attachment Interview has demonstrated similar levels of correspondence between education and RF (Steele & Steele, 2008) and IQ and RF (r values between .27 and .33; compared with .32 in this study; Fonagy et al., 1997). Likewise, IQ and PDI RF have been linked in prior research (Håkansson, Söderström et al. 2018). In this study, the association was most pronounced in the clinical sample (where RF scores were lower) and not significant in the non-clinical group (where there was a broader range of RF scores). In addition, RF scores were significantly lower for mothers with below average levels of nonverbal IQ than they were for mothers with higher levels of IQ. This suggests that while there is some covariance between IQ and RF at the lower levels of IQ, mothers with average or above average nonverbal intelligence can have low or high levels of RF.

Because we excluded mothers with severe learning difficulties in both the clinical and normative samples, we are not able to assess the appropriateness of the PDI RF system for

this population. However, given the moderate association between RF and IQ at the lower end of the nonverbal IQ spectrum, this measure should be used with caution with parents with identified learning problems or documented cognitive limitations. In such cases, behavioural observations may provide a more accurate assessment of the quality of parenting (Shai & Belsky, 2017). This is particularly the case if the parent is also experiencing mental health difficulties, as the association between IQ and RF was most prominent in the clinical sample.

Parental unemployment was also associated with PDI RF, suggesting that this, too, might pose a threat to the discriminant validity of the measure. Mothers who had been unemployed for more than two years had significantly lower levels of RF than those who had been in employment. Once again, there is a question of whether this is a confounding variable or one that we would expect to co-vary with RF. It could be argued that long term unemployment serves as a proxy for sustained poverty, and thus high levels of chronic or "toxic" stress (Shonkoff, 2012). As noted above, stress significantly inhibits mentalization. In addition, mentalization is considered to be an important component to effective social relationships as it makes other people's behavior predictable and understandable (Fonagy, et al., 2002). It is also a crucial element to effective self-regulation within the social environment. The capacity to be offered employment and to remain within a working environment almost always requires effective social skills, and indeed a recent study reported that that level of RF significantly predicted job outcomes (Bly, Wright, & Tuber, 2012). Thus, it is unsurprising that those mothers who had particularly low levels of RF were also more likely to be unemployed. Further, the association may be bidirectional as unemployment may in turn contribute to isolation and lack of opportunities to develop and learn interpersonal abilities and mentalizing. We might therefore consider this as evidence of the concurrent validity of the coding system.

The overall discriminant validity of the PDI RF measure was further confirmed by stepwise linear regressions examining the combined impact of sociodemographic variables -- maternal age, number of children, and education – on PDI RF. Together, these accounted for only a small amount (13%) of the variance in maternal RF ratings. Thus, the sociodemographic characteristics of the sample seem unlikely to play a big part in accounting for RF ratings, indicating the robust discriminant validity of this measure.

Although it is important to identify the demographic variables associated with RF, it appears unlikely that any of those factors identified here pose a threat to the measures' capacity to predict parenting and child outcomes. For example, previous research has shown significant links between fathers' education levels and their AAI-RF scores measured prenatally, but when examined in a hierarchical regression model, it was only the fathers' RF scores that were predictive of the child's attachment to them at 12 months (Steele & Steele, 2008). This suggests that there may be links between some demographic factors and RF, but these associations are modest, theoretically intuitive, and do not necessarily confound the predictive validity of the measure, especially in community samples.

Criterion validity of RF on the PDI

The criterion validity of the coding system was established by examining its capacity to differentiate between levels of risk to the parent-infant relationship. It was hypothesized that the parent-infant dyads in the normative group would be the lowest risk and these mothers would therefore have higher levels of RF than the two high-risk groups. This hypothesis was fully supported; mothers in prison had the lowest levels of RF, mothers with mental health problems had the second lowest levels, and the mothers in the normative group had the highest levels. These findings clearly establish the criterion validity of the PDI in evaluating RF across a range of samples. Other than a recent study linking high psychosocial risk status

26

and low prenatal RF in pregnant woman (Smaling et al., 2015), there have been no studies specifically examining levels of RF across risk groups. However, there are many reasons to link RF and risk status. Maternal mental health problems have been shown in many studies to be related to difficulties within the parent-infant relationship and subsequent infant development (Lyons-Ruth, Connell, & Grunebaum, 1990; Murray, 1992; Rutter, 2005; Sroufe, 2005); thus, the mothers in the clinical group were expected to have lower levels of RF than the mothers in the normative group. Women in prison represent a particularly high-risk group. The prevalence of mental health problems, histories of violence and abuse, substance misuse, and broken attachment relationships for mothers in mother-baby units in prisons are high (Baradon, Fonagy, Bland, Lenard, & Sleed, 2008; Birmingham, Coulson, Mullee, Kamal, & Gregoire, 2006; Black, Payne, Lansdown, & Gregoire, 2004; Byrne & Howells, 2002; Gregoire, Dolan, Birmingham, Mullee, & Coulson, 2010; Sleed, et al., 2013). As all of these factors are considered risks to the evolving parent-infant relationship and to the capacity to mentalize, it was expected that these mothers would have the most difficulty mentalizing about their own and their infant's mental states, as indeed was the case. Thus, researchers using the PDI RF coding system can be confident of its ability to differentiate levels of parental RF across a range of samples.

Strengths and Limitations

There are a number of strengths and limitations to the current study. This is the first study to directly examine and report a number of psychometric properties of a widely used measure. It includes a large and heterogeneous sample of parents and infants from both high-risk and normative populations. Limitations include the fact that the sampling procedure did not capture dyads from either end of the spectrum of risk. In the normative sample, there were relatively high levels of sociodemographic deprivation for many families, such that families

with low risk for relationship disturbances may not have been fully represented. Similarly, the sampling of the high-risk populations did not include dyads with the highest levels of relational risk, such as suspected or known maltreatment. The samples were also not well matched on socio-demographic variables, and only a limited number of demographic variables were captured. In addition, some variables were available for only two of the subsamples. In particular, the measure of nonverbal intelligence was not given to mothers in the prison sample, thus limiting the conclusions that can be drawn in relation to maternal IQ and RF for this particular population.

The findings of this study also highlight some caveats that should be considered when using the coding system. Specifically, ratings of maternal RF may be confounded when the parents have an infant less than two months of age, or if they have below average levels of nonverbal IQ. These findings suggest that the interview and coding system may not be suitable for parents of newborn infants or those with learning difficulties. Furthermore, researchers using the coding system with high-risk samples should be attentive to the distribution of RF scores, as these may be non-normally distributed and may not meet the assumptions of normality that apply to a number of statistical tests.

Conclusion

Parental mentalizing plays a critical role in the social and emotional development of the child. This study marks an important and significant step in developing ways of measuring parental reflective functioning. Our results suggest that the PDI RF coding system is a valid and reliable measure of parental capacities for RF, and that both the interview and the coding system can be used with parents from diverse backgrounds and different levels of risk.

29

References

Arnott, B., & Meins, E. (2007). Links among antenatal attachment representations, postnatal mind-mindedness, and infant attachment security: A preliminary study of mothers and fathers. *Bulletin of the Menninger Clinic, 71*(2), 132-149.

Atlas, J. A. (2001). Review of the Test of Nonverbal Intelligence-Third Edition. In Plake, B.S., & Impara, J.C. (eds). The fourteenth mental measurements yearbook (electronic version) Available from http://search.ebscohost.com/login.aspx?direct=true&db=loh&AN=14072077&site=ehost-live

Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2009). The first 10,000 Adult Attachment Interviews: distributions of adult attachment representations in clinical and non-clinical groups. *Attach Hum Dev, 11*(3), 223-263.

Banks, S. H., & Franzen, M. D. (2010). Concurrent Validity of the TONI-3. *Journal of Psychoeducational Assessment 28*(1), 70-79.

Baradon, T., Fonagy, P., Bland, K., Lenard, K., & Sleed, M. (2008). New Beginnings – an experience-based programme addressing the attachment relationship between mothers and their babies in prisons. *Journal of Child Psychotherapy, 34*(2), 240–258.

Berthelot, N., Ensink, K., Bernazzani, O., Normandin, L., Luyten, P., & Fonagy, P. (2015). Intergenerational transmission of attachment in abused and neglected mothers: the role of trauma-specific reflective functioning. *Infant Mental Health Journal, 36*(2), 200-212.

Birmingham, L., Coulson, D., Mullee, M., Kamal, M., & Gregoire, A. (2006). The mental health of women in prison mother and baby units *The Journal of Forensic Psychiatry & Psychology, 17*(3), 393-404.

Black, D., Payne, H., Lansdown, R., & Gregoire, A. (2004). Babies behind bars revisited. *Archives of Disease in Childhood, 89*(10), 896-898.

Bly, E. M., Wright, A. J., & Tuber, S. B. (2012). Unemployed and poor in New York: The impact of mentalization and Axis II psychopathology on job outcome. *Bulletin of the Menninger Clinic, 76*(2), 101-129.

Bouchard, M.-A., Target, M., Lecours, S., Fonagy, P., & Tremblay, L.-M. (2008). Mentalization in Adult Attachment: Reflective functioning, mental states and affect elaboration compared. *Psychoanalytic Psychology, 25*(1), 47-66.

Brown, L., Sherbenou, R. J., & Johnsen, S. K. (1997). *Test of Nonverbal Intelligence-Third Edition*. Austin, TX: PRO-ED.

Burnett, S., & Blakemore, S. J. (2009). The development of adolescent social cognition. *Annals of the New York Academy of Sciences, 1167*, 51-56. doi:10.1111/j.1749-6632.2009.04509.

Byrne, M. K., & Howells, K. (2002). The Psychological Needs of Women Prisoners: Implications for Rehabilitation and Management. *Psychiatry, Psychology and Law, 9*(1), 34-43.

Chiesa, M., & Fonagy, P. (2014). Reflective function as a mediator between childhood adversity, personality disorder and symptom distress. *Personality and Mental Health, 8*(1), 52-66.

Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition, 107*(2), 705-717.

Ensink, K., Bégin, M., Normandin, L., & Fonagy, P. (2015). Maternal and child reflective functioning in the context of child sexual abuse: Pathways to depression and externalizing difficulties. *European Journal of Psychotraumatology*, *7*(1): 30611.

Ensink, K., Berthelot, N., Bernazzani, O., Normandin, L., & Fonagy, P. (2014). Another step closer to measuring the ghosts in the nursery: Preliminary validation of the Trauma Reflective Functioning Scale. *Frontiers in Psychology, 5*, 1471.

Ensink, K., Fonagy, P., Berthelot, N., Normandin, L., & Bernazzani, O. (2015). Response: Behind the closed doors of mentalizing. A commentary on "Another step closer to measuring the ghosts in the nursery: preliminary validation of the Trauma Reflective Functioning Scale". *Frontiers in Psychology, 6*, 697.

Ensink, K., Leroux, A., Normandin, L., Biberdzic, M., & Fonagy, P. (2015). Assessing reflective parenting in interaction with school-aged children. Submitted for publication.

Ensink, K., Normandin, L., Target, M., Fonagy, P., Sabourin, S., & Berthelot, N. (2015a). Mentalization in children and mothers in the context of trauma: An initial study of the validity of the Child Reflective Functioning Scale. *British Journal of Developmental Psychology, 33*(2), 203-217.

Fischer-Kern, M., Schuster, P., Kapusta, N. D., Tmej, A., Buchheim, A., Rentrop, M., et al. (2010). The relationship between personality organization, reflective functioning, and psychiatric classification in borderline personality disorder. *Psychoanalytic Psychology, 27*(4), 395-409.

Fonagy, P. (2003). The development of psychopathology from infancy to adulthood: the mysterious unfolding of disturbance in time. *Infant Mental Health Journal, 24*(3), 212–239.

Fonagy, P., Gergely, G., Jurist, E., & Target, M. (2002). *Affect Regulation, Mentalization and the Development of the Self*. New York: Other Press.

Fonagy, P., Gergely, G., & Target, M. (2007). The parent-infant dyad and the construction of the subjective self. *Journal of Child Psychology and Psychiatry, 48*, 288-328.

Fonagy, P., Sleed, M., & Baradon, T. (2016). Randomized controlled trial of Parent-Infant Psychotherapy for parents with mental health problems and young infants. *Infant Mental Health Journal, 37*(2), 97-114.

Fonagy, P., Steele, M., Steele, H., Moran, G., & Higgitt, A. (1991). The capacity for understanding mental states: The reflective self in parent and child and its significance for security of attachment. *Infant Mental Health Journal, 12*(3), 201-218.

Fonagy, P., Steele, M., Steele, H., Leigh, T., Kennedy, R., Mattoon, G., & Target, M. (1995). Attachment, the reflective self, and borderline states: The predictive specificity of the Adult Attachment Interview and pathological emotional development. In S. Goldberg, R. Muir, & J. Kerr (Eds.), *Attachment theory: Social, developmental and clinical perspectives* (pp. 233-278). New York: Analytic Press.

Fonagy, P., & Target, M. (1997). Attachment and reflective function: Their role in self-organization. *Development and Psychopathology, 9*, 679-700.

Fonagy, P., Target, M., Steele, H., & Steele, M. (1998). *Reflective-Functioning Manual, version 5.0, for Application to Adult Attachment Interviews*. London: University College London.

George, C., Kaplan, N., & Main, M. (1984). The Adult Attachment Interview. Unpublished manuscript, Department of Psychology, University of California at Berkeley.

Gregoire, A., Dolan, R., Birmingham, L., Mullee, M., & Coulson, D. (2010). The mental health and treatment needs of imprisoned mothers of young children. *Journal of Forensic Psychiatry & Psychology, 21*(3), 378-392.

Grienenberger, J., Kelly, K., & Slade, A. (2005). Maternal reflective functioning, mother-infant affective communication, and infant attachment: Exploring the link between

mental states and observed caregiving behaviour in the intergenerational transmission of attachment. *Attachment and Human Development, 7*(3), 299-311.

Gullestad, F. S., Johansen, M. S., Hoglend, P., Karterud, S., & Wilberg, T. (2013). Mentalization as a moderator of treatment effects: Findings from a randomized clinical trial for personality disorders. *Psychotherapy Research, 23*(6), 674-689.

Ha, C., Sharp, C., Ensink, K., Fonagy, P., & Cirino, P. (2013). The measurement of reflective function in adolescents with and without borderline traits. *Journal of Adolescence, 36*(6), 1215-1223.

Håkansson, U., K. Söderström, et al. (2018). Parental reflective functioning and executive functioning in mothers with substance use disorder. *Attachment and Human Development 20*(2): 181-207.

Katznelson, H. (2014). Reflective functioning: A review *Clinical Psychology Review, 34*(2), 107–117.

Kovács, A. M., Téglás, E., & Endress, A. D. (2010). The Social Sense: Susceptibility to Others' Beliefs in Human Infants and Adults. *Science, 330*(6012), 1830-1834.

Luyten, P., & Fonagy, P. (2015). The neurobiology of mentalizing. *Personality Disorders: Theory, Research, and Treatment, 6*(4), 366-379.

Lyons-Ruth, K., Connell, D. B., & Grunebaum, H. U. (1990). Infants at social risk: Maternal depression and family support services as mediators of infant development and security of attachment. *Child Development, 61*, 85-98.

Mansfield, E. R. and B. P. Helms (1982). Detecting multicollinearity. *The American Statistician, 36*(3a): 158-160.

Murray, L. (1992). The impact of postnatal depression on infant development. *Journal of Child Psychology and Psychiatry, 33*, 543-561.

Newbury-Helps, J., Feigenbaum, J., & Fonagy, P. (2016). Offenders With Antisocial Personality Disorder Display More Impairments in Mentalizing. *Journal of Personality Disorders*, 1-24.

Pajulo, M., Pyykkönen, N., Kalland, M., Sinkkonen, J., Helenius, H., Punamäki, R. L., et al. (2012). Substance-abusing mothers in residential treatment with their babies: importance of pre-and postnatal maternal reflective functioning. *Infant Mental Health Journal, 33*(1), 70-81.

Pedersen, S. H., Lunn, S., Katznelson, H., & Poulsen, S. (2012). Reflective functioning in 70 patients suffering from bulimia nervosa. *European Eating Disorders Review, 20*(4), 303-310.

Rutherford, H. J., Goldberg, B., Luyten, P., Bridgett, D. J., & Mayes, L. C. (2013). Parental reflective functioning is associated with tolerance of infant distress but not general distress: Evidence for a specific relationship using a simulated baby paradigm. *Infant Behavior and Development, 36*(4), 635-641.

Rutter, M. (2005). Environmentally mediated risks for psychopathology: Research strategies and findings. *Journal of the American Academy of Child and Adolescent Psychiatry, 44*(1), 3-18.

Sadler, L. S., Slade, A., Close, N., Webb, D. L., Simpson, T., Fennie, K., et al. (2013). Minding the Baby: Enhancing reflectiveness to improve early health and relationship outcomes in an interdisciplinary home-visiting program. *Infant Mental Health Journal, 34*(5), 391-405.

Schechter, D. S. (2003). Intergenerational communication of maternal violent trauma: Understanding the interplay of reflective functioning and posttraumatic psychopathology. In S. W. Coates & D. S. Schechter (Eds.), *Trauma and Human Bonds* (pp. 115-142). Hillsdale, N.J.: The Analytic Press.

Schechter, D. S., Coates, S. W., Kaminer, T., Coots, T., Zeanah, C. H., Davies, M., et al. (2008). Distorted maternal mental representations and atypical behavior in a clinical sample of violence-exposed mothers and their toddlers. *Journal of Trauma & Dissociation, 9*(2), 123-147.

Schechter, D. S., Coots, T., Zeanah, C. H., Davies, M., Coates, S. W., Trabka, K. A., et al. (2005). Maternal mental representations of the child in an inner-city clinical sample: Violence-related posttraumatic stress and reflective functioning. *Attachment and Human Development, 7*, 313–331.

Schechter, D. S., Myers, M. M., Brunelli, S. A., Coates, S. W., Zeanah, C. H., Davies, M., et al. (2006). Traumatized mothers can change their minds about their toddlers: Understanding how a novel use of videofeedback supports positive change of maternal attributions. *Infant Ment Health Journal, 27*(5), 429-447.

Shai, D., & Belsky, J. (2017). Parental embodied mentalizing: How the nonverbal dance between parents and infants predicts children's socio-emotional functioning. *Attachment and Human Development, 19,* 191-219.

Sharp, C., & Fonagy, P. (2008). The parent's capacity to treat the child as a psychological agent: Constructs, measures and implications for developmental psychopathology. *Social Development, 17*(3), 737-754.

Sharp, C., Venta, A., Vanwoerden, S., Schramm, A., Ha, C., Newlin, E., et al. (2016). First empirical evaluation of the link between attachment, social cognition and borderline features in adolescents. *Compr Psychiatry, 64*, 4-11.

Shonkoff, J.P. (2012). Leveraging the biology of adversity to address the roots of disparities in health and development. *Proceedings of the National Academy of Sciences in the United States of America, 109* (Suppl. 2), 17302-17307.

Slade, A. (2005). Parental reflective functioning: An introduction. *Attachment and Human Development, 7*(3), 269-281.

Slade, A., Aber, J. L., Bresgi, I., Berger, B., & Kaplan, M. (2004). The Parent Development Interview – Revised. Unpublished protocol. New York, NY: The City University of New York.

Slade, A., Bernbach, E., Grienenberger, J., Levy, D., & Locker, A. (2004). Addendum to Fonagy, Target, Steele, & Steele reflective functioning scoring manual for use with the Parent Development Interview. Unpublished Manuscript. New York: The City College and Graduate Center of the City University of New York.

Slade, A., Grienenberger, J., Bernbach, E., Levy, D., & Locker, A. (2005). Maternal reflective functioning, attachment, and the transmission gap: A preliminary study. *Attachment and Human Development, 7*(3), 283-298.

Sleed, M., Baradon, T., & Fonagy, P. (2013). New Beginnings for mothers and babies in prison: A cluster randomized controlled trial. *Attachment and Human Development 15*, 349-367.

Smaling, H., Huijbregts, S., Suurland, J., Van Der Heijden, K., Van Goozen, S., & Swaab, H. (2015). Prenatal reflective functioning in primiparous women with a high-risk profile. *Infant Mental Health Journal, 36*(3), 251-261.

Sroufe, L. A. (2005). Attachment and development: A prospective, longitudinal study from birth to adulthood. *Attachment and Human Development, 7*(4), 349-367.

Stacks, A. M., Muzik, M., Wong, K., Beeghly, M., Huth-Bocks, A., Irwin, J. L., et al. (2014). Maternal reflective functioning among mothers with childhood maltreatment histories: Links to sensitive parenting and infant attachment security. *Attachment and Human Development, 16*(5), 515-533.

Steele, H. and M. Steele (2011). On the Origins of Reflective Functioning. In F. N. Busch (Ed.). *Mentalization: Theoretical considerations, research findings, and clinical implications.* New York, Taylor & Francis. 133-158.

Stern, D. N. (1995). *The Motherhood Constellation: A Unified View of Parent-Infant Psychotherapy.* New York, N.Y.: BasicBooks.

Suchman, N., DeCoste, C., Castiglioni, N., McMahon, T. J., Rounsaville, B., & Mayes, L. (2010). The Mothers and Toddlers Program, an attachment-based parenting intervention for substance using women: Post-treatment results from a randomized clinical pilot. *Attachment & Human Development, 12*(5), 483-504.

Suchman, N., Decoste, C., McMahon, T. J., Rounsaville, B., & Mayes, L. (2011). The mothers and toddlers program, an attachment-based parenting intervention for substance-using women: Results at 6-week follow-up in a randomized clinical pilot. *Infant Mental Health Journal, 32*(4), 427-449.

Suchman, N., Legow, N., Decoste, C., & Castiglioni, N. (2008). The Mothers and Toddlers Program: Preliminary findings from an attachment-based parenting intervention for substance-abusing mothers. *Psychoanalytic Inquiry, 25*(3), 499–517.

Verhage, M. L., Schuengel, C., Madigan, S., Fearon, R. M., Oosterman, M., Cassibba, R., et al. (2016). Narrowing the transmission gap: A synthesis of three decades of research on intergenerational transmission of attachment. *Psychological Bulletin, 142*(4), 337-366.

Vermote, R., Lowyck, B., Luyten, P., Vertommen, H., Corveleyn, J., Verhaest, Y., et al. (2010). Process and outcome in psychodynamic hospitalization-based treatment for patients with a personality disorder. *Journal of Nervous and Mental Disease, 198*(2), 110-115.

Winnicott, D. W. (1956). Primary maternal preoccupation. In D.W.Winnicott (Ed.), *Collected papers: Through paediatrics to psycho-analysis* (pp. 300-305). London: Tavistock, 1958.

Table 1. Description of sample

	Normative N = 57	Clinical N = 121	Prison N = 163	Total N = 341
Child gender: N (%)*				
Female	29 (51%)	53 (44%)	99 (61%)	180 (53%)
Male	28 (49%)	68 (56%)	64 (39%)	159 (47%)
Child age in months: *				
Mean (SD)	7.2 (2.6)	4.1 (3.1)	4.6 (4.5)	4.9 (3.9)
Range	0.9–12.6	0.5–12.3	0.1–23.6	0.1–23.6
Mother age in years: *				
Mean (SD)	33 (5)	31 (6)	27 (6)	30 (6)
Range	21-50	19-42	18-42	18-50
Number of other children: N (%)				
None (first time mothers)	42 (75%)	75 (62%)	84 (53%)	201 (60%)
One	10 (18%)	28 (23%)	39 (25%)	77 (23%)
Two	3 (5%)	14 (12%)	19 (12%	36 (11%)
Three or more	1 (2%)	4 (3%)	17 (11%)	22 (6%)
Mother ethnicity: N (%)*				
White	42 (74%)	74 (61%)	89 (55%)	205 (60%)
Black	2 (4%)	17 (14%)	52 (32%)	71 (21%)
Asian	6 (11%)	14 (12%)	10 (6%)	30 (9%)
Mixed	5 (9%)	8 (7%)	11 (7%)	24 (7%)
Other	2 (4%)	8 (7%)	1 (0.5%)	11 (3%)
Mother education*				
Did not complete high school	1 (2%)	6 (5%)	63 (41%)	70 (21%)
Basic (high school equivalent)	10 (18%)	48 (40%)	42 (28%)	100 (30%)
Further (vocational training)	5 (9%)	14 (12%)	38 (25%)	57 (17%)
Higher (degree or higher)	41 (72%)	53 (44%)	9 (6%)	103 (31%)
Estimated Nonverbal IQ* **(TONI-3)**				
Mean (SD)	111 (11)	105 (3)	Not measured	107 (12)
Range	83-145	72-140		72-145

* Significant differences between groups (see text)

Table 2. Social Exclusion Criteria

	Normative N=57	Clinical N=121	Total N=178
Social Exclusion Criteria: N (%)			
Eligible for income support*	10 (18%)	61 (50%)	47 (27%)
Long term unemployed*	6 (11%)	41 (34%)	47 (27%)
In temporary/overcrowded accommodation*	9 (16%)	38 (31%)	47 (27%)
Mother is single or unpartnered*	3 (5%)	42 (35%)	45 (25%)
Mother has chronic illness or disability	12 (21%	21 (17%)	33 (19%)
Mother has history of foster or institutional care	0	3 (3%)	3 (2%)
Socially isolated (recent relocation)	14 (25%)	46 (38%)	60 (34%)
Mother under 20 years of age	0	2 (2%)	2 (1%)
Mother has previous psychiatric diagnosis*	10 (18%)	80 (66%)	90 (51%)

* Significant differences between groups

Table 3. Reflective Functioning Scores for each demand question of the PDI

	Normative N = 56	Clinical N = 118	Prison N = 149	Total N = 323
Clicked				
Mean (SD)	4.6 (1.7)	3.9 (1.5)	3.8 (1.4)	4.0 (1.5)
Not clicked:				
Mean (SD)	4.9 (1.6)	4.3 (1.7)	2.8 (2.0)	3.7 (2.0)
Relationship affecting personality				
Mean (SD)	4.1 (1.4)	3.3 (1.5)	3.0 (1.5)	3.3 (1.5)
Joy				
Mean (SD)	3.3 (1.3)	3.0 (1.3)	2.7 (1.3)	2.9 (1.3)
Pain or difficulty				
Mean (SD)	3.7 (1.3)	3.2 (1.4)	2.9 (1.5)	3.1 (1.4)
Having child changed you				
Mean (SD)	3.2 (1.4)	3.3 (1.4)	2.9 (1.5)	3.1 (1.4)
Angry				
Mean (SD)	4.3 (2.0)	4.4 (1.6)	3.0 (2.0)	3.9 (2.0)
Guilty				
Mean (SD)	4.3 (1.9)	4.1 (1.8)	3.4 (1.9)	3.8 (1.9)
Needy				
Mean (SD)	3.9 (1.8)	3.9 (1.6)	Not asked	3.9 (1.4)
Child upset				
Mean (SD)	5.0 (1.5)	4.3 (1.6)	3.6 (1.7)	4.1 (1.7)
Rejected				
Mean (SD)	3.6 (2.0)	2.8 (1.8)	2.1 (1.6)	2.6 (1.8)
Parents				
Mean (SD)	5.4 (2.0)	4.8 (1.9)	3.6 (1.8)	4.4 (2.0)
Child's feelings about separation				
Mean (SD)	3.6 (1.6)	3.1 (1.4)	3.1 (1.5)	3.2 (1.5)
Mother's feelings about separation				
Mean (SD)	4.0 (1.4)	3.6 (1.4)	3.5 (1.3)	3.6 (1.4)
Losing				
Mean (SD)	2.8 (1.8)	2.6 (1.6)	Not asked	2.7 (1.7)
Overall RF Score				
Mean (SD)	4.6 (1.4)	4.1 (1.4)	3.4 (1.4)	3.8 (1.5)

Table 4. Correlations between PDI demand question RF scores and overall RF score

	Pearson's correlation with Overall RF score
Clicked	.68**
Not clicked	.66**
Relationship affecting personality	.66**
Joy	.55**
Pain or difficulty	.63**
Having child changed you	.63**
Angry	.67**
Guilty	.65**
Needy	.60**
Child upset	.70**
Rejected	.51**
Parents	.75**
Child's feelings about separation	.57**
Mother's feelings about separation	.56**
Losing	.39**

** All correlations are significant at the p < .001 level (2-tailed)

Table 5. Correlation between overall RF score and socio-demographic variables

	Normative N = 56	Clinical N = 118	Prison N = 149	Total N = 323
Child gender	-.01	.01	-.06	.02
Child age	.06	-.14	.23**	.14*
Mother age	.10	.13	.09	.22***
First time mother	.23	.39***	-.11	.17**
Mother educated past high school	.24	.13	.21*	.27**
Maternal Nonverbal IQ	.17	.35**		.32***
Social exclusion criteria:				
Eligible for income support	-.17	-.24**		-.27***
Long term unemployed	-.33*	-.32***		-.35***
In temporary/overcrowded accommodation	-.03	.01		-.03
Mother is single or unpartnered	-.09	-.18		-.21**
Mother has chronic illness or disability	.17	-.05		.03
Mother has history of foster institutional care	n/a	-.09		-.09
Socially isolated (recent relocation)	-.01	.03		-.01
Mother under 20 years of age	n/a	-.01		-.02
Mother has previous psychiatric diagnosis	.12	.01		-.05

* Correlation is significant at the 0.050 level (2-tailed)

** Correlation is significant at the 0.010 level (2-tailed)

*** Correlation is significant at the 0.001 level (2-tailed)

Table 6. Regression equations of demographic predictors of maternal RF

Predictor variables	Beta (SE)	β	t	p
Constant	2.49 (.39)		6.42	< .001
Mother's age	.05 (.01)	.19	3.30	< .001
Number of other children the mother has	-.34 (.08)	-.23	-4.14	< .001
Mother received post-high school education	.54 (.17)	.18	3.21	< .001

Table 7. Correlations between predictor variables

	Mother's age	Child age	Mother received post-high school education	Number of other children the mother has
Mother's age	1.00	.14*	.33***	.18**
Child age		1.00	.08	.02
Mother received post-high school education			1.00	-.15**
Number of other children the mother has				1.00

* Correlation is significant at the 0.050 level (2-tailed)

** Correlation is significant at the 0.010 level (2-tailed)

*** Correlation is significant at the 0.001 level (2-tailed)