Maternal reflective functioning and precursors of theory of mind and executive functioning in 20-month-old children

Inge Kremer – Berg-Andersen

1076477

Leiden University

Faculty of Social and Behavioral Sciences

Developmental Psychopathology in Education and Child Studies

Research Master Thesis, June 2015

Supervisor: Dr. Stephan C.J. Huijbregts

Second reader: Prof. Dr. Hanna Swaab

Abstract

Objective To examine (a) the effect of maternal reflective functioning (RF) on precursors of theory of mind (ToM) and executive functioning (EF) in 20-month-old children, (b) the relation between ToM and EF at 20 months, and (c) the effectiveness of an early intervention program aimed, among others, at improving maternal RF. **Method** The sample consisted of 118 mother-child dyads. Maternal RF was assessed during pregnancy using the Pregnancy Interview-Revised and at 20 months using the Parent Development Interview. At 20 months children's ToM understanding was examined using a simple visual perspective (VP) taking task, a discrepant desires (DD) task and an imitation task; EF performance was assessed using a delay task (inhibition) and a working memory (WM) task called 'hide the pots'. The Child Behavior Checklist was conducted to assess children's problem behavior. **Results** Children of mothers low on maternal RF, especially child-related RF, showed significantly worse VP-taking capacities compared with children of mothers average/high on RF. A significant moderating effect of children's problem behavior was present; children with high attention problems or a high withdrawn level were more affected by their mothers' low RF capacities. No unequivocal and significant effects of maternal RF on DD, imitation and EF performance were found. In addition, only a significant correlation between DD and WM performance was present. The effectiveness of the early intervention program regarding improving RF could not be confirmed. **Conclusions** Early intervention programs should focus on improving child-related aspects of maternal RF of mothers at risk for being low on RF, especially when their children show problem behavior. Future research should reexamine both the effect of maternal RF on ToM and EF, and the relationship between ToM and EF at e.g. 2.5 years.

Keywords: maternal reflective functioning, theory of mind, executive functioning, visual perspective taking, intervention program, problem behavior, temperament

Introduction

In the last few decades there has been an explosion of research into the development of theory of mind (ToM; understanding other people's thoughts, feelings, desires and intentions; Baron-Cohen, Tager-Flusberg, & Cohen, 1993) and executive functioning (EF; cognitive processes that regulate behavior; Geurts & Huizinga, 2011) of typically developing children and of clinical groups (Baron-Cohen, 1995; Ozonoff, Pennington & Rogers, 1991). Hughes and colleagues have been pioneers in showing a ToM-deficit in children with externalizing behavioral disorders, more specifically conduct disorder (CD) and oppositional defiant disorder (ODD; Happé & Frith, 1996; Hughes & Ensor, 2008). Barkley (1997) introduced a cognitive model indicating that inhibitory dysfunction, leading to poor EF in general, is the core deficit in attention deficit hyperactivity disorder (ADHD). This line of thought was confirmed by subsequent research linking executive dysfunction to ADHD (e.g. Happé, Booth, Charlton & Hughes, 2006). Furthermore, a high comorbidity (30-50%) between CD/ODD and ADHD has been reported (Spencer, 2006). Research has shown that the development of externalizing disorders and the underlying neurocognitive functions of ToM and EF involve a gene-environment interaction (Hughes et al., 2005; Hughes & Ensor, 2008; Swaab, Bouma, Hendriksen & König, 2011), and that quality of parenting matters with respect to the development of EF (Hammond, Müller, Carpendale, Bibok, & Liebermann-Finestone, 2012; Bernier, Carlson & Whipple, 2010; Mileva-Seitz, 2015), and of ToM (Meins et al., 2002; Meins, Fernyhough, Arnott, Leekam & De Rosnay, 2013; Ruffman, Slade, Devitt & Crowe, 2006). Furthermore, externalizing disorders, as well as the underlying neurocognitive dysfunctions, are associated with negative long-term outcomes such as school dropout, substance abuse, criminality and incarceration, and this link is more profound for children with early-onset behavioral problems (Moffitt et al., 2011; Sharp, Fonagy & Goodyer, 2008). Given these outcomes, studying the effect of quality of parenting on the early development of ToM and EF is crucial in order to develop effective prevention and intervention programs for mother-child dyads at risk.

The role of parenting; maternal reflective functioning

Research and interventions on quality of parenting at first focused on techniques for obtaining and maintaining discipline, however, the focus has shifted towards a parent's ability to treat the child as a psychological agent (a person that can reason about its own and other people's explicit intentions, goals and beliefs), also called ToM or parental mentalization (Baron-Cohen et al., 1993). Research has shown that the neural base of mentalizing consists of regions including the superior temporal sulcus, the adjacent temporo-parietal junction, the temporal poles and the medial prefrontal cortex (Frith & Frith, 2003; Frith & Frith, 2006). Interestingly, Winnicott's concept of 'good enough parenting' also seems to apply to the accuracy of parental mentalization; mothers low on accuracy had children with poorer psychosocial adjustment, no significant differences were found between the average and high maternal accuracy groups (Sharp, Fonagy & Goodyer, 2006). Several constructs have been developed to operationalize this parental mentalizing ability, among which parent reflective functioning (RF) and maternal mind mindedness (MMM) are the ones most frequently used in research and interventions (Sharp & Fonagy, 2008). In the present study mothers' mentalization capacity is operationalized by maternal RF. Slade (2005) defined RF as "the essential human capacity to understand behavior in light of underlying mental states and intentions". Maternal RF can be defined as the mother's ability to reflect upon her own and the child's internal mental states and use this capacity to guide her responses to her child (Fonagy, Gergely, Jurist & Target, 2002). So, RF involves the mother's expressions based on metacognitive representations about herself and her relationship with her child, whereas MMM, a measure of mothers' mental state talk, examines the observed online, real life interactions between parent and child (Sharp &

Fonagy, 2008). Research has shown that maternal RF is related to quality of parenting (Grienenberger, Kelly & Slade, 2005), attachment security (Grienenberger et al., 2005; Slade, Grienenberger, Bernbach, Levy & Locker, 2005), and positive child outcomes (Benbassat & Priel, 2012). A study by Smaling and colleagues (2015) indicated that mothers with a high-risk status, as assessed using the Mini-International Neuropsychiatric Interview-plus, demonstrated significantly lower RF capacities compared with low-risk mothers. Furthermore, the level of maternal education, substance use during pregnancy, and size of social support network showed to be the strongest predictors of prenatally measured RF for mothers at risk (Smaling et al., 2015). Notwithstanding the above, more research is needed, especially to explore the associations between parental RF and child development.

Interventions aimed at improving reflective functioning

A wide range of mentalization-based therapies have been developed to enhance mentalizing abilities in adults and children. Slade describes two reflective parenting programs, (a) a group intervention for parents of infants, toddlers and preschoolers called Parents First, and (b) a preventive program developed for high risk first time pregnant young women called Minding the Baby (Slade, 2010). The latter intervention consists of weekly home visits by a pediatric nurse practitioner and a clinical social worker starting prenatally until the children are two years old. This program was developed to enhance parental RF and attachment security, next to maternal physical and mental health (Slade, 2010). Findings from the Minding the Baby pilot study indicated less disorganized attached children and more securely attached children as well as less rapid second births and less child maltreatment in the participating group versus the control group (Ordway et al., 2014). A randomized trial study examining the effects of the Minding the Baby intervention showed that participating mothers described their children as having less externalizing problems at one to three years post intervention compared to the control group, however the results didn't show the expected increase in parental RF for the participating versus the control group (Ordway et al., 2014). It is clear that these promising though not satisfying results need further research.

Development of ToM abilities

ToM or mentalizing can be defined as the capacity to attribute thoughts, emotions and intentions to other people (Baron-Cohen et al., 1993). The mentalizing ability makes it possible to explain and predict behavior, and is usually observed in false belief tasks that require children to imagine someone else's thoughts. Different levels of mentalizing are distinguished; first order mentalizing (emerging at three-to-four years), second order mentalizing (present from five to six years of age), and higher order mentalizing, that develops into late adolescence. Even though research suggests an innate preference of babies for social stimuli (a few weeks old babies smile more and vocalize more towards humans than towards objects) evidence of mentalizing becomes apparent from approximately 18 months (Frith & Frith, 2003). More precisely, early aspects of ToM are thought to develop in toddlerhood and be implicit features manifested in behavior rather than explicit language (Laranjo, Bernier, Meins & Carlson, 2010). One such feature is the appreciation of other people's visual perception. Research has shown that visual perspective taking develops in two steps. At first, children acquire an understanding that others need to have their eyes open and directed toward an object, without something blocking their vision, in order to be able to see an object (Laranjo et al, 2010). Research has shown that infants between 12.5 and 18 months of age already behave according to this simple (level 1) perspective taking (Luo & Baillargeon, 2007; Poulin-Dubois, Sodian, Metz, Tilden & Schoeppnes, 2007). According to the simulation theory (Harris, 1992), in the second step, children come to understand that others may see a different appearance of an object if they look at it from another position (level 2 visual perspective taking). This ability has proven to be developmentally and conceptually related to false belief understanding, both emerging around three years of age (Moll & Meltzoff, 2011). Another early feature of ToM development is the ability to understand others' desires and act accordingly, an aspect that is thought to be acquired by 18 months (Repacholi & Gopnik, 1997). This age is also significant for the intended reliable imitation of actions by the child (Meltzoff, 1995), and for the onset of pretend play, as infants from this age laugh and don't get confused when their mothers pretend to use a banana as a telephone indicating their understanding of pretence (Leslie, 1987).

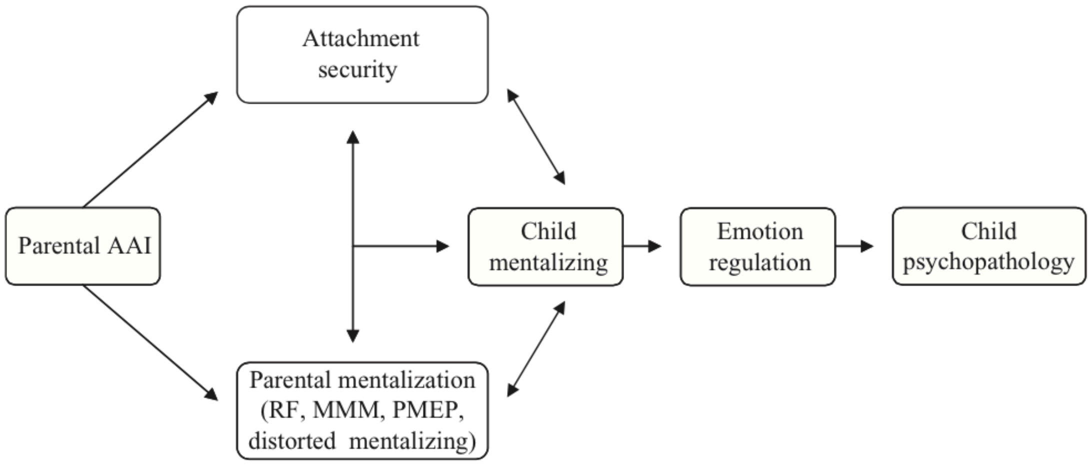

Figure 1. Sharp & Fonagy's testable model for the development of child psychopathology through mentalization (reproduced from Sharp & Fonagy, 2008)

A longitudinal twin study with a large sample of 60 months olds showed that 44% of the variation in ToM scores could be explained by nonshared environmental factors and 41% by shared environmental influences, leaving common genetic factors to account for only 15% (Hughes et al., 2005). In this respect Sharp and Fonagy (2008) introduced a model linking parental mentalization capacities (as measured among others by RF) to child mentalizing abilities (figure 1): parental attachment representations (as measured with the Adult Attachment Interview (AAI)) lead to attachment security and adequate parental mentalization, both of which influence each other and affect child mentalizing abilities; which in turn is linked with emotion regulation and finally child psychopathology. Research has supported the suggested relationship between parent attachment representations, parental mentalization and attachment security (Arnott & Meins, 2007; Fonagy, Steele, Moran, Steele & Higgitt, 1991), and the expected function of attachment security (Fonagy, 2003; Fonagy & Target, 2005). The proposed effect of maternal mentalization on child mentalizing has been sustained by several findings using MMM to operationalize maternal mentalization. Early MMM has been found to significantly explain children's ToM performance at 45, 48 and 51 months (Meins et al., 2002; Meins et al., 2013; Ruffman et al., 2006). A study assessing the early manifestations of children's ToM showed that MMM at 12 months was positively related to understanding of discrepant desires and level 1 visual perspective taking at 26 months of age (Laranjo et al., 2010). Moreover, a follow-up study indicated that MMM at 12 months predicted understanding of false belief and level 2 visual perspective taking at 49 months over and above perspective taking at 26 months (Laranjo, Bernier, Meins & Carlson, 2014). However, as far as we know no research has been conducted examining the relation between maternal RF and manifestations of ToM at 20 months.

Development of EF: inhibition and working memory

Executive functions (EF) are cognitive processes (inhibition, working memory, cognitive flexibility and planning) that regulate behavior in such a way that behavior can be efficient and goal directed. Especially in new and unfamiliar situations that require a flexible behavioral approach EF are crucial. Therefore, EF play an important role in the understanding of social situations and social behavior (Swaab et al., 2011). As the development of inhibition and working memory (WM) forms the hart of EF and starts at an early age compared to the development of cognitive flexibility and planning, we will focus our account on inhibition and WM.

Inhibitory control is known to be very difficult for young children (Diamond, 2013). Infants of six to 11 months have a strong tendency to directly reach for and grasp a visible object they desire (Diamond, 2013). However, 18 month olds seem to be able to exercise inhibitory control, at which age individual differences in inhibiting abilities become apparent (Rothhart, Derryberry & Posner, 1994). Nevertheless, the development of inhibitory control continues into childhood and even adolescence (Davidson, Amso, Anderson & Diamond, 2006). Research indicates that good inhibitors at four years of age more likely possess better inhibitory abilities at 14 years (Eigsti et al., 2006). Furthermore, a significant role of inhibition in the development of children's emotion regulation capacities and social development has been found (Hirshfeld-Becker et al., 2003; Kochanska, Murray & Harlan, 2000).

Working memory (WM) is crucial for making sense of the world around us, as understanding the world requires holding in mind earlier events and relating them to new developments. Besides, WM is necessary for doing math, turning instructions into actions, creativity, reasoning, planning and decision-making. Furthermore, it should be noted that WM supports inhibitory control and vice versa (Diamond, 2013). WM and short-term memory are closely related, as WM requires holding information in mind. However, WM and short-term memory are two distinct constructs; they are linked to different neural subsystems and follow a different development path (Diamond, 2013). Research using Piaget's human cognitive development 'A not B test' showed that infants as young as 7.5 to 12 months can uncover a toy hidden in one of two possible locations when there is no delay, but make perseveration errors when a delay of one to five seconds is introduced (Diamond & Goldman-Rakic, 1989). From 12 months onwards infants succeed on the A not B test even with delays as long as ten seconds, indicating the emergence of the ability of holding information in mind (Diamond & Goldman-Rakic, 1989). However, the ability to hold more information in mind or to mentally manipulate stored information takes much longer to develop (Davidson et al. 2006).

Research shows that quality of parenting, operationalized e.g. by parental scaffolding (Hammond et al., 2012), autonomy support, maternal sensitivity and maternal mind-mindedness (Bernier et al., 2010), and parental sensitivity (Mileva-Seitz, 2015) matters in the development of EF in young children. Moreover, Blair, Raver and Berry (2014) found evidence that in addition to higher quality parenting (operationalized as parent responsiveness and sensitivity) enhancing EF development, children's level of EF capacities predicted positive change in parenting quality. Research examining the role of parenting regarding inhibition, specifically, showed that a positive parenting style, higher monitoring and lower discipline, inconsistency and negative controlling are associated with proper inhibition development (Roskam et al., 2014). Furthermore, research suggests an effect of quality of parenting on WM development; e.g. a longitudinal study showed that high levels of maternal sensitivity, as observed during a disciplinary task at three years, were associated with lower WM problem scores at four years, as measured with the Behavior Rating Inventory of Executive Function-Preschool Version (BRIEF-P; Kok et al., 2014); a study of Bernier and colleagues (2014) found significant correlations between MMM and maternal autonomy support, measured at 12-15 months and WM scores at 18 months of age. However, as far as we know no studies have examined the role of parenting, operationalized by maternal RF, on child inhibition and WM.

Gender, temperament and language

During (early) childhood, girls perform better than boys regarding both EF, especially inhibition (Berlin & Bohlin, 2002; Carlson & Moses, 2001; Kochanska et al., 2000; Mileva-Seitz et al., 2015) and ToM (e.g. Calero, Salles, Semelman & Sigman, 2013; Walker, 2005). Furthermore, gender seems to be a moderator regarding the relationship between parenting and child outcomes, and in general it seems that boys are more vulnerable to quality of parenting. For example, McFadyen-Ketchum, Bates, Dodge and Pettit (1996) found that high maternal coercion and lack of affection were associated with the development of aggression in boys; no such results were found with respect to girls. The results of a study by Griffin, Botvin, Scheier, Diaz and Miller (2000) indicated that more parental monitoring was associated with less alcohol use and less delinquency, in boys only. Mileva-Seitz and colleagues (2015) found that parental sensitivity was associated with better attention, however this association only was present regarding boys. Besides, boys also seem to benefit more from a positive home environment (Horton, Kahn, Perera, Barr & Rauh, 2012).

The vast majority of prior research indicates direct linear effects of child temperament on social competence, including emotion regulation and inhibitory control (e.g. Rothbart & Ahadi, 1994; Rothbart & Jones, 1998). Nevertheless, in line with a vulnerability or predisposition model, it is thought that early temperament may predispose a child to certain outcomes, with other (external) processes also playing a significant role in predicting adverse outcomes (Rothbart & Ahadi, 1994). For example, findings showed that a parenting intervention was successful for mothers with highly negative infants, but not for mothers with infants low on negativity (Anzman-Frasca, Stifter, Paul & Birch, 2014). Another study indicated temperament as a moderator of the effects of parental depression on child behavior problems (Jessee, Mangelsdorf & Wong, 2012). Blair, Denham, Kochanoff & Whipple (2004) found that both temperament and emotion regulation predict the quality of children's social functioning, however interaction effects between temperament and emotion regulation seem to predict social functioning more accurately.

Research suggests a role of child language abilities regarding the relationship between parenting and ToM and EF development; a link between parent mentalization abilities and child language acquisition through attachment security has been found

(Fonagy, 2003), plus an association between language capacities and child ToM and EF abilities (Sharp, Fonagy & Goodyer, 2008). Furthermore, the 2013' study by Meins and colleagues pointed to a significant role for receptive verbal ability; a significant model showed both a direct and indirect link between MMM and ToM, the latter being mediated through child receptive verbal ability.

The relationship between the development of ToM and EF

In their review on the development of ToM and executive control Perner and Lang (1999) conclude that research with three to five year old children clearly shows a developmental link between improved EF (e.g. self-control) and ToM development. Later research has confirmed this association; the relation persists when controlling for age differences between children and verbal ability (Carlson & Moses, 2001; Hughes & Ensor, 2007), and similar associations have been found in multiple cultures (Chasiotis, Kiessling, Hofer & Campos, 2006). However, researchers have struggled to interpret these relatively robust findings. One possible explanation is that ToM and EF tasks both demand common EF skills, though the majority of research does not support this account (Bull, Phillips & Conway, 2008; Carlson, Claxton & Moses, 2015; Perner, Lang & Kloo, 2002). A functional interdependence of EF and ToM is a second possible interpretation, with most research indicating that executive skills enable the acquisition of mental-state concepts (Carlson et al., 2015; Dennis, Agostino, Roncadin & Levin, 2009; Hughes & Ensor, 2007). Moreover, research suggests that some parameters of EF play a more important role: more profound associations have been found using conflict or cognitive inhibition (flexibly inhibiting and activating competing cognitive responses) as opposed to delay inhibition (stopping or delaying a response) and WM (Carlson et al., 2015; Carlson & Moses, 2001; Dennis et al., 2009). The proposed explanation is that conflict inhibition tasks involve both inhibition and working memory (Carlson et al., 2015; Carlson & Moses, 2001; Hughes & Ensor, 2007). The majority of research has used samples of children aged three years or older to assess the association between EF and ToM. Using a sample of two year olds (*M* = 2.37 years, *SD* = 4 months), Hughes & Ensor (2005) found that EF and ToM were significantly related with a medium to large effect size. However, a study with a sample consisting of 24 and 39 months old children only showed a significant association between EF and ToM capacities of the 39 months olds (Carlson, Mandell & Williams, 2004).

The current study

The current study focuses on maternal RF as a measure of mentalization, both prenatal and postnatal, and its possible influence on children's ToM and EF at 20 months of age. Based on the majority of literature, it is hypothesized that 20-month-old children of mothers with low RF capacities show less early ToM and EF abilities compared with children of mothers average or high on RF. Furthermore, the role of child language, gender and temperament are assessed. We expect to find that child language abilities positively predict ToM and EF performance, and that girls demonstrate better ToM and EF abilities than boys. Moreover, significant moderating effects of both child gender and child temperament regarding the effect of maternal RF on ToM and EF are expected. It is hypothesized that boys and children with a more 'difficult' temperament are more vulnerable for mothers low on maternal RF, and therefore show lower ToM and EF capacities compared with these children of mothers average or high on RF. Next, it is assessed whether the often found correlation between child EF and ToM development is also present in the current sample of under two year olds. In addition, the current study examines the proposed effect of an existing early intervention program aimed, among others, at improving maternal RF of high-risk mothers. It is hypothesized that RF skills of mothers who participated in the program have increased significantly prenatally to postnatally, and this improvement is significantly larger compared to a high-risk control group, and a low-risk control group. The implications of the findings for both theory and practice are explored.

Method

Participants and procedure

The present study is part of a large longitudinal study in the Netherlands, called the Mother-Infant Neurodevelopment Study in Leiden (MINDS-Leiden), following mothers and their first-born infants from 27 weeks of pregnancy till the child is 2.5 years old. The aim of the larger study is to examine neurobiological and neurocognitive processes that have been related to early problem behavior. The larger study was approved by the ethics committee of the Department of Education and Child Studies of the Faculty of Social Behavioral Sciences at Leiden University and by the Medical Research Ethics Committee at Leiden University Medical Centre. Written informed consent was acquired from all participating women. Pregnant women were recruited via pregnancy fairs, midwifery clinics, hospitals and prenatal classes. To participate in the study women had to be between 17 and 25 years of age during the pregnancy, be expecting their first baby and be sufficiently fluent in the Dutch language. Participants who turned out to have a severe drug addiction, severe psychiatric problems or an IQ below 70 were excluded from the study. The larger study consists of five measurement moments: a home visit when mothers are in their third trimester of pregnancy, ideally around 27 weeks of pregnancy (T1), a second home visit when infants are approximately 6 months (T2), a laboratory measurement at 12 months (T3), a third home visit at 20 months (T4), and a second laboratory measurement when children are 2.5 years old (T5). Based on an intake-screening interview at T1 mothers were assigned to either a high-risk (HR) group or a low-risk control group (LR-CG). For a HR classification the following criteria were used: (a) current psychiatric disorder(s), or substance use during pregnancy, or (b) presence of at least two of the following: limited social support network, no achieved secondary education, unemployment, financial problems or poverty, single mother and mothers' age below 20. Subsequently, HR mothers were randomly assigned either to an intervention group (HR-IG), receiving coaching until the children were 2.5 years old, or a care-as-usual (control) group (HR-CG). At T4 early aspects of ToM were assessed using a task that taps into level 1 visual perspective taking, a task examining understanding of discrepant desires, and a task measuring imitation performance. Furthermore, when children were 20 months old early EF was examined using a task to measure WM abilities and a delay task to assess response inhibition capacities.

For participants to be included in the present study, data regarding the Pregnancy Interview-Revised (PI-R) and Parent Development Interview (PDI; conducted at T1 and T4 respectively) had to be present. Subsequently, the final sample of the present study consisted of 118 Dutch mothers and their first-born 20-month-old children. Participants were mainly Caucasian (82.2%) and the most frequent reported highest level of achieved education was secondary education. Mothers' mean age at T1 was 22.23 (*SD* = 2.37). At T4 the children's mean age was 20.4 months (64 boys).

Measurement instruments

Maternal RF. Maternal RF was assessed by trained interviewers during two home-visits; at T1 prenatal RF was examined using a Dutch translation (Smaling & Suurland, 2011) of the PI-R (Slade, Patterson & Miller, 2007), and at T4 a Dutch translation (Smaling, 2013) of the PDI (Slade, Bernbach, Grienenberger, Levy & Locker, 2005) was administered. Both instruments are semi structured clinical interviews that take about 45 minutes to administer and were digitally recorded and subsequently transcribed verbatim. The PI-R consists of 22 questions intended to examine the mother's experiences of her pregnancy, and her expectations and fantasies about the relationship to come with her unborn baby. The PDI contains 45 questions aimed at eliciting parents' representation of themselves as parents, of their children and of their relationships with their children. The original Reflective Functioning manual was used for scoring the PI-R and PDI (Fonagy et al., 1998), next to the Addendum to the Reflective Functioning Scoring Manual for use of the PI (Slade et al., 2007) and the Addendum to Reflective Functioning Scoring Manual for use with the PDI (Slade et al., 2005). Regarding both the PI-R and PDI the extent of RF is coded on a continuum from low to high reflective abilities, ranging from -1 (negative RF) to 9 (full or exceptional RF). According to the manuals scores < 5 indicate negative, absent or low RF, whereas scores of > = 5 represent evidence of RF (Slade et al., 2007; Slade et al, 2005). Regarding the PDI, next to the total RF score, additional sub scores are obtained for mothers' self-related RF (representations of themselves as parents) and child related RF (representations of their children and of their relationships with their children). Analyses were performed using the total score of prenatal RF (PI-R), the total score of postnatal RF (PDI), the total score of child-related RF (PDI) and the total score of self-related RF (PDI). Inter-rater reliability regarding the total RF score and individual passage scores were .90 and .87 respectively for the PI-R (based on 15% double coded interviews) and .93 and .80 respectively for the PDI (based on 15 interviews coded by a second rater).

RF Intervention. The HR-IG participated in a coaching intervention consisting of home visits by trained therapists; weekly visits in the first year, starting in the third trimester of pregnancy, changing to visits every two weeks and finally monthly visits till the child's second birthday. The program aimed at developing participants' mother role by increasing their RF capacities, next to improving mothers' life style and social network.

Visual Perspectives. An adaptation of the visual perspectives task, originally developed by Carlson, Mandell and Williams (2004) was used to assess children's early understanding of visual perspective. The examiner showed the child a newly introduced toy from a box (Sesame Street Bert, a book, a rubber ducky, maracas, and a snow dome with Miffy) and asked the child to show the toy to his/her mother who was sitting a few feet away. To be able to show the toy appropriately the child had to perform a physical or vocal act as the mother (1) had her eyes closed, (2) covered her eyes with her hands, (3) had her eyes blindfolded, (4) was sitting with her back towards the child, or (5) the toy needed special pointing (Miffy). For each of the five trials the child could receive a score varying from 0 to 5 (0 = the child doesn't react to the request and/or doesn't show interest in the toy, 1 = the child doesn't show the toy to mother or drops it close to her, 2 = the child holds the toy close to mother, but doesn't perform the necessary correction so the mother can't see the toy, 3 = the child partly performs a correction, but breaks off the correction before mother can see the toy, 4 = the child performs the correction but doesn't show the toy subsequently, 5 = the child performs the necessary correction and subsequently shows the toy to mother), adding up to a total score of 0 to 25.

Discrepant Desires. The discrepant desires task that was conducted, is based on the discrepant desires task used by Carlson and colleagues (2004), and is an adaption of the food-request procedure (Repacholi & Gopnik, 1997). The task aimed to assess children's capacity to understand that people can have other desires than their own. The examiner offered the child two different snacks in small pieces in two bowls placed on a tray and asked the child to choose one. The experimenter then took the tray back and acted like she disgusted the chosen snack and loved the other snack, which act was repeated. While providing again the two snacks on the tray the examiner asked the child if she could have some. After the child's response the tray was removed for a moment and then shoved into reachable position of the child again while the experimenter again asked if she could have some. After the second child's response, the tray was removed, after which the experimenter repeated the disgusted/loving act twice. Two additional trails took place following the same procedure as before. Scores consisted of the number of correct responses: 1 = the child gives the examiner the snack the examiner likes; 0 =

the child gives the snack the examiner disgusts or doesn't give a snack at all, providing a total score of 0 to 4.

ADOS Imitation task. Functional and symbolic imitation was measured using an imitation task of the Autism Diagnostic Observation Schedule (ADOS; Lord, et al., 1989; Lord et al, 2000), a semi structured observation instrument aimed at diagnosing and assessing autism spectrum disorders. The examiner introduced a toy by naming it and showing a physical and vocal act with the toy, after which the child was told it's his/her turn (e.g., drinking gestures and noises were made by the examiner when introducing a cup). The task consisted of two practice trials, three functional imitations, and three symbolic imitations (the examiner shows a cube when saying: "Now this is a cup", followed by the drinking gestures and noises). Scores consisted of the number of correct functional and symbolic imitation trials (0-6). Research examining the psychometric properties of the ADOS–Generic indicated a high inter-rater reliability of the functional and symbolic imitation task (Lord et al., 2000).

Delay task. A delay task was used to measure children's early inhibition capacities based on Kochanska and colleagues' (1996) snack delay task; children needed to withhold a prepotent response to grab or touch an attractive toy (a colorful magic wand) placed in front of them right after the examiner had told them not to touch the toy for a moment. Inhibition was coded on a continuum from low to high inhibitory control, ranging from 1 (the child touches the toy before the examiner places it on the table) to 9 (the child doesn't touch the toy during the 30 seconds test). Besides, a simple pass/fail scoring was obtained, failing meaning any form of touching or playing with the magic wand during 30 seconds.

Hide the Pots. Hide the Pots is a downward adaptation of the Spin the Pots task (Hughes & Ensor, 2005), and was developed by Bernier and colleagues (2010) to test WM in children between 18-24 months of age. An attractive toy (Sesame Street Ernie) was hidden underneath one of three cups of different colors, and the child was subsequently asked: "Where is Ernie?" The task consisted of three practice trials without a delay, followed by three test trials with a two seconds delay (placing a box over the three cups). Scores consisted of the number of correct test trials (0-3).

ECBQ. A short version (excluding the Extraversion/Surgency scale) of the Early Childhood Behavior Questionnaire (ECBQ; Putnam, Gartstein, Rothbart, 2006) was filled out by mothers at T4. This 80 items parent report is designed to measure toddler temperament. Items are rated on a 7-point Likert-style scale ranging from 'never' to 'always', plus a non-applicable option. Putnam and colleagues (2006) demonstrated a moderate inter-rater reliability for most of the 18 scales and an adequate internal consistency for all scales. Furthermore, their research revealed a three-factor structure (negative affectivity, effortful control and surgency), similar to the Childhood Behavior Questionnaire (CBQ; Rothbart, Ahadi, Hershey & Fisher, 2001) and the Infant Behavior Questionnaire-Revised (IBQ-R; Gartstein & Rothbart, 2003) and showed longitudinal stability relations. Factor scores varied from 0 to 7 with higher scores indicating higher levels of negative affectivity and better effortful control.

CBCL & PASEC. At T4 mothers filled out a combination questionnaire of the Dutch translation of the Child Behavior Checklist for 1½ to 5-year old children (CBCL; Achenbach & Rescorla, 2000), a widely used parent report to identify problem behavior in children, and the Dutch version of the Physical Aggression Scale for Early Childhood (PASEC; Alink et al., 2006) to measure the level of children's physical aggression. The 99 items of the CBCL and the 11 items of the PASEC were scored on a 3-point Likert scale, indicating the extent to which the item is applicable to the child according to the mother (0 = not at all; 1 = a bit; 2 = certainly or often). Total aggression scores of the PASEC vary from 0 to 22, a higher score indicating more aggressive behavior. Items all elicit explicit aggressive behavior, e.g. item 1: "Is cruel for animals", item 4: "Fights a lot" and item 10: "Threatens others to hit them". Internal consistency of the Dutch version of the PASEC can be considered good (Alink et al., 2006). With respect to the CBCL, we incorporated both the factor Internalizing problems (scores: 0-36), and Externalizing problems (scores: 0-46) for the present study. Furthermore, two of the four scales of the internalizing factor were used: (1) the Emotionally Reactive scale (scores varying from 0–18; 0-5: normal range; 6-8: borderline range; >= 9 clinical range), and the Withdrawn scale (range of scores: 0–16; 0-4: normal range; 5: borderline score; >= 6: clinical range), and both scales of the externalizing factor: (1) the Attention Problems scale (scores: 0-10; 0-5: normal range; 6: borderline score; >= 7: clinical range), and the Aggressive Behavior scale (scores: 0-36; 0-20: normal range; 21-24: borderline range; >= 25: clinical range). The aggressive behavior scale of the CBCL contains some of the items of the PASEC (eliciting explicit aggressive behavior), plus items that express more general problem behavior, e.g. item 8 "Can't stand waiting, everything has to happen immediately", item 15: "Provoking" and item 20: "Disobedient". Research has confirmed the seven-syndrome model of the CBCL in various societies, including the Netherlands (Ivanova et al., 2010), and indicates good psychometric properties of the CBCL (Achenbach & Rescorla, 2000).

Reynell & Schlichting. The first 11 items of the adapted Dutch version of the Reynell Developmental Language Scale (Schaerlaekens, 1995) were used to measure child language reception at T4, adding up to a total score of 0 to 11. During the same home visit the children's active language capacities were examined using the first 12 items of the Schlichting test of language production, resulting in a total score of 0 to 12.

Data analysis

The Statistical Package for the Social Sciences (IBM SPSS; Version 21.0) was used performing all statistical analyses. Prior to testing the main hypotheses, data were inspected thoroughly with respect to missing values, outliers and violations of assumptions applying to the statistical tests used. An outlier was defined as a score deviating at least three standard deviations from the mean. Outliers were imputed to the closest most extreme score within three standard deviations from the mean. No missing values were present in the maternal RF data. Regarding the other data the multiple imputation method was used to impute the missing values. Given the large sample size, the assumption of normality was mainly checked visually by inspecting the histograms and Q-Q-plots. Variables were log transformed if distributions were positively skewed. Composite scores were created only when variables significantly correlated and the correlation could be considered at least moderate ($r > .4$). First, we examined the expected effect of maternal RF on early ToM and EF capacities. Correlational analyses between maternal RF on the one hand and the individual measures of ToM and EF on the other hand were performed. Subsequently, we created a group of mothers low on RF and a group average or high on RF. Analyses of variance, chi-square test of the independence of two categorical variables and logistic regression were used to assess whether the two RF groups differed regarding children's early manifestations of ToM and EF. If applicable, child temperament was used as a covariate. Furthermore, the possible main effect of gender and the possible moderating roles of child gender and temperament were examined. In this respect groups of children with an easy versus a difficult temperament were created. The four subgroups of easy vs. difficult temperament and A/H vs. low maternal RF each had to contain a minimum of three participants to allow further moderation analysis. Subsequently, the role of child language abilities regarding ToM and EF performance was assessed, using regression analysis and analysis of variance. Second, we examined the relationship between early ToM and EF capacities by performing correlational analyses. Last, the effect of the coaching program was examined. After we assessed whether maternal RF increased from T1 to T4, using repeated measures analysis of variance, and whether the coaching vs. non-coaching group differed in level of RF, we examined the effect of the coaching program using mixed design analysis of variance.

Results

Preliminary analyses

Maternal RF. Parameters of the four measures of maternal RF are depicted in table 1: prenatal RF, postnatal RF, child-related (postnatal) RF and self-related (postnatal) RF. As the total scores of prenatal and postnatal scores were significantly and moderately correlated (r = .42, p < .001), a composite score of RF was created computing the mean of prenatal and postnatal maternal RF (table 1). Further analyses were performed using the five measures of maternal RF. Correlations of these five measures are depicted in table 2. As research suggests the principle of good enough parenting also applies to maternal RF (Sharp, Fonagy & Goodyer, 2006), for each of the maternal RF measures, a group low on maternal RF (low RF group) and a group average/high on maternal RF (A/H RF group) was created. Although the PI-R and PDI manuals indicate that scores < 5 indicate negative, absent or low RF, whereas scores of >= 5 represent evidence of RF (Slade et al., 2007; Slade et al, 2005), our low RF group only included scores that could be considered low with respect to the current sample. As the mean score of each RF measure equaled approximately 4, and both the median and mode equaled 4, we chose our low RF group to include scores of 2 to 3 and the A/H RF group to include scores of 3.5 and higher (table 1).

RF intervention. The three experimental groups differed significantly in mean maternal age (F (2, 115) = 21,52, p < .001, η_p^2 = .27), with the HR-IG having the lowest mean age (M = 20.19, SD = 2.42, N = 21), followed by the HR-CG (M = 21.57, SD = 2.32, N = 37), and the LR-CG (M = 23.35, SD = 1.69, N = 60).

Visual perspectives (VP). Actual scores on the visual perspective task covered the full range of possible scores (0 – 25; M = 15.11; SD = 5.21).

Discrepant desires (DD). Actual scores on the discrepant desires task ranged from 0 to 4, the full range of possible scores, however the majority of children (69.5%) received a score of 0 (M = 0.60, SD = 1.10). Given these results a simple binary variable was computed, indicating that children either failed (score of 0) or passed the task (a score of 1-4; 30.5% of the children).

ADOS imitation task. Actual scores for the functional and symbolic imitation task consisted of the total range of possible scores (0-6, M = 3.41, SD = 2.05).

Table 1. *Main parameters of maternal RF*

Maternal RF measurement	*N*	*M*	*Median*	*Mode*	*Min*	*Max*	*SD*
Prenatal total RF (PI)	118	3.80	4	4	2	6	.94
Low	*48*	*2.85*	*3*	*3*	*2*	*3*	*.36*
A/H	*70*	*4.44*	*4*	*4*	*4*	*6*	*.61*
Postnatal total RF (PDI)	118	4.25	4	4	2	8	1.00
Low	*26*	*2.96*	*3*	*3*	*2*	*3*	*.20*
A/H	*92*	*4.61*	*4*	*4*	*4*	*8*	*.83*
Child-related RF (PDI)	118	4.10	4	4	2	7	.84
Low	*26*	*2.96*	*3*	*3*	*2*	*3*	*.20*
A/H	*92*	*4.42*	*4*	*4*	*4*	*7*	*.65*
Self-related RF (PDI)	118	3.94	4	4	2	7	.97
Low	*39*	*2.87*	*3*	*3*	*2*	*3*	*.34*
A/H	*79*	*4.47*	*4*	*4*	*4*	*7*	*.71*
Composite score RF	118	4.02	4	4	2	6.5	.82
Low	*22*	*2.91*	*3*	*3*	*2*	*3*	*.25*
A/H	*96*	*4.28*	*4*	*4*	*3.5*	*6.5*	*.68*

Note: A/H = Average/High; composite score RF = mean of pre- and postnatal RF

Table 2. *Correlations maternal RF (N = 118)*

	Prental RF	Postnatal RF	Child-related RF	Self-related RF	Composite score RF
Prenatal RF	1	.42 ***	.35 ***	.26 **	.83 ***
Postnatal RF		1	.72 ***	.69 ***	.85 ***
Child-related (postnatal) RF			1	.48 ***	.64 ***
Self-related (postnatal) RF				1	.57 ***
Composite score RF					1

Note: ** is significant at $p < .01$; *** is significant at $p < .001$.

Delay task. Inspection of the actual continuum of scores, ranging from 1 to 9, showed a non-normal distribution (M = 5.05, SD = 3.45); most 20-month-olds either touched the toy even before the examiner placed it on the table (score = 1; 22.9%) or didn't touch the toy at all (score = 9; 36.4%). Therefore, the simple binary pass or fail (37.3% vs. 62.7%) score was used in further analyses.

Hide the Pots. The actual scores consisted of the full range of possible scores (0-3; M = 2.03, SD = 0.86), however, only 5 out of the 118 participants answered all three trials incorrectly. As the possibility of passing one out of three trials by chance equals 1, we recoded the variable as follow: a score of 0 or 1 = 1 (N = 32); 2 = 2 (N = 46) and 3 = 3 (N = 40). Further analyses were performed using these three categories.

ECBQ, PASEC, CBCL. The main parameters of the temperament and problem behavior scales can be found in table 3. As the distribution of Aggressive Behavior as measured with the PASEC, the factor Internalizing Problems, and the scales Emotionally Reactive and Withdrawn of the CBCL were positively skewed, values were log-transformed for further analyses. Furthermore, a group of children low on aggressive behavior (PASEC) and a group high on aggressive behavior were created (table 3). Regarding the four scales of the CBCL a group of participants with scores in the normal range (normal group) and a borderline/clinical group (BC group) were created according to the manual of the CBCL (table 3).

Reynell & Schlichting. The Reynell task for language comprehension turned out to be less difficult for the 20-month-old children than the Schlichting test for language production. The distribution of scores of the Reynell test was negatively skewed with scores varying from 2 to the maximum score of 11 (M = 8.74, SD = 2.19), whereas a positively skewed distribution was found regarding the scores of the Schlichting test (M = 3.03, SD = 2.62). A significant correlation of the Reynell and Schlichting was found, that can be considered large (r = .59, p < .001). Given these results a composite score of child language capacities was computed adding the language comprehension and language production scores. The composite language score was normally distributed with scores ranging from 2 to 21 (M = 11.76, SD = 4.28).

Table 3. *Main parameters of child temperament measurements*

Child temparement	*N*	*M*	*SD*	Actual range of scores
Effortful Control (ECBQ)	118	4.56	.52	3.31 – 6.03
Negative Affectivity (ECBQ)	118	2.94	.45	2.00 - 4.14
Aggressive Behavior (PASEC)	118	3.21	2.70	0 – 12
Low group	*103*	*2.44*	*1.70*	*0 – 5*
High group	*15*	*8.53*	*2.29*	*6 - 12*
Emotional Reactivity (CBCL)	118	2.80	1.89	0 – 8
Low group	*103*	*2.24*	*1.26*	*0 – 5*
BC group	*15*	*6.60*	*.74*	*6 – 7*
Withdrawn (CBCL)	118	1.78	1.46	0 – 6
Low group	*108*	*1.43*	*.91*	*0 – 4*
BC group	*10*	*5.60*	*.52*	*5 – 6*
Attention Problems (CBCL)	118	4.56	1.94	0 – 10
Low group	*85*	*3.65*	*1.31*	*0 -5*
BC group	*33*	*6.91*	*1.16*	*6 - 10*
Aggressive Behavior (CBCL)	118	13.64	4.83	1 – 26
Low group	*106*	*12.58*	*3.78*	*1 - 20*
BC group	*12*	*23.08*	*1.88*	*21 - 26*
Internalizing problems	118	8.52	4.72	1 – 28
Externalizing problems	118	18.21	6.03	3 – 33

Note: BC Group = Group with the respective temperament aspect in the Borderline-Clinical range

Maternal RF and child language capacities. Before we examined the expected effect of maternal RF on children's performance on the ToM and EF tasks, the role of child language, gender and child temperament regarding maternal RF was assessed, as the results of these analyses indicated possible covariates and/or moderators regarding the effect of maternal RF on Tom and EF. Children's language abilities were not significantly correlated with maternal RF (using any of the five RF measurements). Nevertheless, analyses of variance showed that mothers with a low composite RF score seem to have children with less language capacities (F (1, 116) = 5.52, p < .05, η_p^2 = .05; table 4) compared with children of mothers in the A/H group. A trend in the same direction was found regarding both postnatal RF (F (1, 116) = 2.95, p < 0.1, η_p^2 = .02) and child-related RF (F (1, 116) = 3.73, p < 0.1, η_p^2 = .03; table 4). The results indicated neither significant differences nor trends in language capacities between the groups when examining prenatal RF and self-related RF (table 4). Given these findings, children's language score was not used as a covariate in the main analysis regarding the effect of maternal RF on child ToM and EF, however the effect of child language capacities on ToM and EF performance was assessed separately.

Table 4. *Mothers low vs. A/H on maternal RF and children's language capacities*

Maternal RF measurement	$Group_{RF}$	*N*	*M*	*SD*	*F*	*p-value*	η_p^2
Prenatal total RF (PI)	Low	48	11.17	4.62	1.57	n.s.	-
	A/H	70	12.17	4.02			
Postnatal total RF (PDI)	Low	26	10.50	4.59	2.95	< .10	.02
	A/H	92	12.12	4.15			
Child-related RF (PDI)	Low	26	10.35	4.88	3.73	< .10	.03
	A/H	92	12.16	4.04			
Self-related RF (PDI)	Low	39	11.28	4.57	0.73	n.s.	-
	A/H	79	12.00	4.14			
Composite score RF	Low	22	9.86	4.92	5.52	< .05	.05
	A/H	96	12.20	4.03			

Maternal RF and child gender. Analyses of variance indicated no significant differences between boys and girls regarding their mothers' level of maternal RF, when using any of the five maternal RF measurements.

Maternal RF and child temperament. The relationship between maternal RF and child temperament was examined using all five measures of RF and the nine measures of child temperament. Results of the correlational analyses are depicted in table 5. Mothers higher on self-related RF reported their children to have a more difficult temperament; a higher negative affectivity ($r = .20$, $p < .05$), more externalizing problems ($r = .23$, $p < .05$), more attention problems ($r = .21$, $p < .05$), more aggressive behavior, as measured with the CBCL ($r = .20$, $p < .05$), and higher emotional reactivity ($r = .20$, $p < .05$). Furthermore, mothers lower on prenatal RF and composite RF reported their children as behaving more aggressive, when measured with the PASEC ($r = -.20$, $p < .05$, and $r = -.16$, $p = .08$ respectively). Besides, mothers with higher levels of child-related RF reported their children as more emotionally reactive ($r = .19$, $p < .05$), however less withdrawn ($r = -.19$, $p < .05$). Analyses of variance showed results in line with the correlational analysis. Children of mothers low on prenatal RF were reported to show more aggressive behavior (PASEC; *M* = 3.88, *SD* = 2.65), compared with children of mothers A/H on prenatal RF (*M* = 2.76, *SD* = 2.67; $F(1, 116) = 6.16$, $p < .05$, $\eta_p^2 = .05$). Next, our results showed that mothers low on self-related RF seem to report their children to have an easier temperament compared to mothers A/H on self-related RF; with significantly less externalizing problems (M = 16.49, SD = 5.85 versus M = 19.06, SD = 5.97; $F(1, 116) = 4.92$, $p < .05$, $\eta_p^2 = .04$), to behave significantly less aggressive (*M* = 12.31, *SD* = 4.46 versus *M* = 14.30, *SD* = 4.89; $F(1, 116) = 4.60$, $p < .05$, $\eta_p^2 = .04$), as less negative affective (*M* = 2.83, *SD* = .49 versus *M* = 3.00, *SD* = .42; $F(1, 116) = 3.59$, $p = .06$, $\eta_p^2 = .03$), and higher on effortful control (*M* = 4.69, *SD* = .46 versus *M* = 4.50, *SD* = .54; $F(1, 116) = 3.67$, $p = .06$, $\eta_p^2 = .03$), when compared to mothers A/H on self-related RF. No other significant differences between mothers low versus A/H on RF were found regarding child temperament.

Table 5. *Correlations maternal RF and child temperament (N = 118)*

Child temperament	Prental RF	Postnatal RF	Child-related RF	Self-related RF	Composite score RF
Negative affectivity (ECBQ)	-.04	.09	.07	.20 *	.04
Effortful control (ECBQ)	-.01	-.06	.10	-.13	-.04
Aggressiveness (PASEC)	-.20 *	-.08	-.12	.02	-.16 #
Internalizing problems (CBCL)	.02	-.01	.02	.07	.00
Emotional reactivity (CBCL)	.05	.15	.19 *	.20 *	.12
Withdrawn level (CBCL)	-.07	-.11	-.19 *	-.01	-.11
Externalizing problems (CBCL)	-.02	.07	-.12	.23 *	.03
Attention problems (CBCL)	-.02	.13	.09	.21 *	.07
Aggressive behavior (CBCL)	-.02	.03	.06	.20 *	.00

Note: * is significant at $p < .05$; # is significant at $p < .10$.

Main analyses

Maternal RF predicting child ToM and EF performance.

The effect of maternal RF on VP-taking. A trend was found regarding the correlation between maternal prenatal RF and child VP-taking (r = .14, p = .07). Neither significant correlations nor trends were found regarding the other measures of maternal RF and VP-taking. Further analyses using analysis of variance showed that 20-month-old children with mothers in the low RF group, as measured with the composite RF score and the child-related RF score, performed significantly worse on the VP-taking task than children of mothers in the A/H RF group (F (1, 116) = 5.22, p < .05, η_p^2 = .04; and F (1, 116) = 5.92, p < .05, η_p^2 = .05 respectively; table 6). A trend in the same direction was found regarding prenatal RF; mothers low on prenatal RF seemed to have children that had less VP-taking capacities at 20 months than children of mothers in the A/H group (F (1, 116) = 2.95, p < 0.1, η_p^2 = .02; table 6). No significant differences were found between the two groups with respect to total postnatal RF and self-related RF (table 6).

Table 6. *Mothers low vs. average/high on maternal RF and their children's VP-taking*

Maternal RF measurement	$Group_{RF}$	*N*	*M*	*SD*	*F*	*p-value*	η_p^2
Prenatal total RF (PI)	Low	48	14.13	5.80	2.95	< .10	.02
	A/H	70	15.79	4.68			
Postnatal total RF (PDI)	Low	26	14.15	6.32	1.13	n.s.	-
	A/H	92	15.38	4.85			
Child-related RF (PDI)	Low	26	12.96	5.84	5.92	< .05	.05
	A/H	92	15.72	4.88			
Self-related RF (PDI)	Low	39	14.77	5.64	0.25	n.s.	-
	A/H	79	15.28	5.01			
Composite score RF	Low	22	12.86	6.38	5.22	< .05	.04
	A/H	96	15.63	4.79			

Child gender. Factorial analyses of variance neither indicated a main effect of child gender on child VP-taking nor a moderating role for child gender regarding the effect of maternal RF on child VP-taking; non-significant results were found using all five measures of maternal RF.

The role of child temperament. Correlational analysis showed that VP-taking performance was significantly related to children's withdrawn level (r = -.22, p < .05), and total internalizing problems (r = -.21, p < .05). No significant correlations were found regarding the other measurements of temperament and VP-taking. Given these results and the preliminary data analysis, indicating differences between the RF groups with respect to children's withdrawn level, children's internalizing problems was used as possible covariate in the analysis to reduce within-group error variance. When using child-related RF and composite RF, the results indicated children's internalizing problems as a significant covariate (F (1, 115) = 4.37, p < .05, *partial* η_p^2 = .04, and F (1, 115) = 4.15, p < .05, *partial* η_p^2 = .04, respectively), next to a significant effect of maternal RF on VP-taking capacities after controlling for the effect of children's internalizing problems (F (1, 115) = 5.06, p < .05, *partial* η_p^2 = .04, and F (1, 115) = 4.15, p < .05, *partial* η_p^2 = .04, respectively). Furthermore, the results showed that the effect of prenatal RF remained nearly significant (F (1, 115) = 2.91, p < .1, *partial* η_p^2 = .03), and children's internalizing problems was found to be a significant covariate (F (1, 115) = 5.17, p < .05, *partial* η_p^2 = .04). Neither trends nor significant effects were found regarding postnatal RF and self-related RF as predictor of VP taking when children's internalizing problems was used as covariate in the analysis. Next to these analyses of covariance, we examined a possible moderating role of child temperament, as literature indicates a temperament vulnerability model (Rothbart & Ahadi, 1994). Factorial analysis of variance indicated that the effect of RF level on VP-taking performance was different for children with a difficult temperament as measured with the withdrawn and attention problems scales of the CBCL. Children's withdrawn level showed to be a

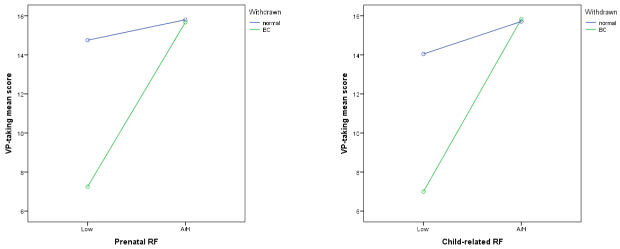

Figure 2. Interactions of withdrawn level (normal vs. BC) and maternal RF (low vs. A/H)

Table 7. *Children's withdrawn level and VP-taking*

Maternal RF measurement		*Withdrawn*	M_{RF}	SD_{RF}	*N*
Prenatal RF	Low	Normal	14.75	5.29	44
		B/C	7.25	7.63	4
	A/H	Normal	15.80	4.76	64
		B/C	15.67	4.13	6
Child-related RF	Low	Normal	14.05	5.03	22
		B/C	7.00	7.17	4
	A/H	Normal	15.71	4.94	86
		B/C	15.83	4.21	6

significant moderator of the relationship between maternal RF level and child VP-taking capacities. Using prenatal RF, there was a significant main effect of prenatal RF group on VP-taking (F (1, 114) = 7.77, p < .01, η_p^2 = .06) and of withdrawn level (F (1, 114) = 5.05, p < .05, η_p^2 = .04). Besides, a significant interaction effect of RF group x withdrawn group was found (F (1, 114) = 4.71, p < .05, η_p^2 = .04). Similar results were found using child-related RF; there was a significant main effect of RF group (F (1, 114) = 9.31, p < .01, η_p^2 = .08), a significant main effect of withdrawn level (F (1, 114) = 4.05, p < .05, η_p^2 = .03), and a significant interaction effect of RF group and withdrawn level (F (1, 114) = 4.34, p < .05, η_p^2 = .04). More specifically these findings show that highly withdrawn children from mothers low on maternal RF show significantly worse VP-taking capacities at 20 months than highly withdrawn children from mothers of A/H RF group; no such differences were found regarding children with normal withdrawn levels (figure 2). However, it should be noted that these findings are based on very low numbers of participants falling into the highly withdrawn groups (table 7). Moderation analyses using postnatal, self-related and composite RF were not pursued, as the groups of highly withdrawn children with mothers low on maternal RF were too small (N = 1 or 2). Further analysis indicated child attention problems as a significant moderator regarding the effect of prenatal RF on VP-taking capacities. Although results did not show a significant main effect of attention problems (F (1, 114) = 1.09, p = n.s), a significant main effect of prenatal RF group was found (F (1, 114) = 5.99, $p < .05$, $\eta_p^2 = .05$). Moreover, there was a significant interaction between the level of prenatal RF and the level of attention problems (F (1, 114) = 4.28, $p < .05$, $\eta_p^2 = .04$), indicating that children with low versus high attention problems were affected differently by their mothers' level of prenatal RF. Children with BC attention problems of mothers low on prenatal RF performed significantly worse on the VP-taking task than children with these attention problems but mothers in the A/H RF group; no such differences between children from mothers low versus A/H on prenatal RF were found regarding children with normal levels of attention problems (figure 3; table 8).

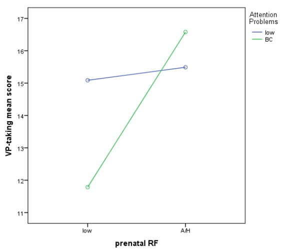

Figure 3. Interaction of level of attention problems (normal vs. BC) and maternal RF (low vs.A/H)

Table 8. *Children's attention problems and VP-taking*

Maternal RF measurement		*Attention problems*	M_{RF}	SD_{RF}	N
Prenatal RF	Low	Normal	15.09	5.62	34
		B/C	11.79	5.77	14
	A/H	Normal	15.49	4.74	51
		B/C	16.58	4.54	19

The effect of child language on VP-taking. Correlational analysis showed that VP-taking performance was significantly related to child language capacities (r = .31, p < .001). As preliminary data analysis had shown that child language capacities differ among children of mothers low vs. A/H on maternal RF, separate regression analysis was performed to examine the effect of language. Child language capacities had a significant effect on VP-taking performance (B = .38, t = 3.52, p = .001, $R^2_{adjusted}$ = .09), indicating that children with better comprehensive and productive language capacities performed better on the VP-taking task.

Table 9. *Possible predictors of DD performance (N = 118; Df = 1)*

Measurement	Group	*N*	% passed on DD task	χ^2	*p-value*
Prenatal total RF (PI)	Low	48	25%	1.16	n.s.
	A/H	70	34%		
Postnatal total RF (PDI)	Low	26	23%	.87	n.s.
	A/H	92	33%		
Child-related RF (PDI)	Low	26	19%	2.00	n.s.
	A/H	92	34%		
Self-related RF (PDI)	Low	39	26%	.65	n.s.
	A/H	79	33%		
Composite score RF	Low	22	23%	.77	n.s.
	A/H	96	32%		
Gender	Boys	64	28%	.38	n.s.
	Girls	54	33%		

Predicting DD performance; maternal RF. Chi-square tests of the independence of two categorical variables showed no significant association between DD performance and RF group, as measured with the five maternal RF measurements (table 9). Furthermore, no significant association was found between gender and the success rate of DD performance (table 9). Data exploration showed a possible positive relation between child language capacities and DD success. Subsequently, binary logistic regression analysis indicated that children's language capacities significantly predict DD success (Wald χ^2 (1, N = 118) = 5.38, p < .05; B = .12, SE = .05; *Nagelkerke's* R^2 = .07). When a child's composite language scoring increases 1 point, the odds for passing the DD task becomes 1.12 times as high. The results did not show significant effects of child temperament on DD performance, although a trend was found regarding aggressive behavior (PASEC); successful children on the DD task seemed to be less aggressive (M = .47, SD = .30) than children who failed (M = .57, SD = .28; F (1, 117) = 2.82, p < .1).

Table 10. *Possible predictor of children's imitations capacities (Df (1, 116).*

Measurement	Group	*N*	$M_{imitation}$	$SD_{imitation}$	*F*	*p-value*
Prenatal total RF (PI)	Low	48	3.06	2.06	2.30	n.s.
	A/H	70	3.64	2.03		
Postnatal total RF (PDI)	Low	26	3.38	2.28	.00	n.s.
	A/H	92	3.41	2.00		
Child-related RF (PDI)	Low	26	2.96	2.22	1.58	n.s.
	A/H	92	3.53	2.00		
Self-related RF (PDI)	Low	39	3.44	1.98	.01	n.s.
	A/H	79	3.39	2.10		
Composite score RF	Low	22	3.18	2.04	.32	n.s.
	A/H	96	3.46	2.06		
Gender	Boy	64	3.41	2.03	.00	n.s.
	Girl	54	3.41	2.10		

Predicting child imitation capacities; maternal RF. Neither trends nor significant correlations between maternal RF and children's imitation capacities were present. Moreover, no significant differences in imitation performance were found between children of mothers in the low RF group compared with children of mothers with A/H RF capacities (table 10), plus child gender and temperament were no significant moderators in this respect. Further analysis did not show a significant main effect of gender on the imitation capacities of 20-month-old children (table 10). However, results indicated that child language abilities significantly related to children's imitation performance (r = .24, p < .01). Moreover, significant correlations were found between imitation performance on the one hand and children's withdrawn level ($r = -.22$, $p < .05$) and total internalizing problems ($r = -.23$, $p < .05$) on the other hand. The other measurements of child temperament did not show any significant correlations. Hierarchical multiple regression analysis revealed a best fitting model ($F (2, 115) = 4.97$, $p < .05$, $R^2_{adjusted} = .08$) with child language capacities significantly positively predicting VP-taking ($B = .10$, $t = 2.38$, $p < .05$) and children's internalizing problems significantly negatively predicting VP-taking performance ($B = -1.86$, $t = -2.23$, $p < .05$).

Predicting child inhibition; maternal RF. Pearson's Chi-square test indicated that maternal RF level was not significantly associated with performance on the inhibition delay task (table 11). Further analysis showed that gender and inhibition were significantly associated; 28% of boys were successful at the delay task versus 48% of the girls ($\chi^2 (1, N = 118) = 5.02, p < .05$). If the child happened to be a girl the odds of passing the delay task was 2.38 times higher than if the child was a boy. Neither significant effects of child language abilities nor child temperament on children's inhibition performance were found.

Table 11. *Possible predictors of children's inhibition capacities (N = 118; Df = 1))*

Maternal RF measurement	$Group_{RF}$	*N*	% passed the delay task	χ^2	*p-value*
Prenatal total RF (PI)	Low	48	42%	.66	n.s.
	A/H	70	34%		
Postnatal total RF (PDI)	Low	26	42%	.36	n.s.
	A/H	92	36%		
Child-related RF (PDI)	Low	26	35%	.10	n.s.
	A/H	92	38%		
Self-related RF (PDI)	Low	39	44%	.99	n.s.
	A/H	79	34%		
Composite score RF	Low	22	36%	.92	n.s.
	A/H	96	38%		
Gender	Boy	64	28%	5.02	< .05
	Girl	54	48%		

Predicting child WM capacities; maternal RF. Data exploration showed possible relationships between WM performance on the one hand and postnatal RF, child-related RF, child language abilities and child withdrawn level on the other hand. No such relationships were found regarding child gender and the other measurements of maternal RF and child temperament. However, as preliminary analysis showed a (nearly) significant relationship between postnatal RF and child-related RF on the one hand and child language abilities and child withdrawn level on the other hand, subsequent multinominal logistic regression was performed using only the maternal RF measures as predictors. Results showed that mothers' level of child-related RF (low vs. A/H) significantly predicted WM score. The odds of a child with a mother low on child-related RF obtaining a score of 2 instead of 1 on the WM task is 8.79 times higher than for a child with a mother A/H on child-related RF (Wald χ^2 (1, N = 118) = 7.54, p < .01; B = 2.17, SE = .79; *Nagelkerke's* R^2 = .11). Furthermore, the odds for this child to receive a 2 score instead of 3 is 2.8 times higher than for a child with a mother A/H on child related RF (Wald χ^2 (1, N = 118) = 3.88, p < .05; B = 1.02, SE = .52; *Nagelkerke's* R^2 = .11). However, no such differences were found regarding obtaining a WM score of 1 versus 3 (table 12). Similar results were obtained using postnatal RF as predictor; low postnatal RF significantly increased the odds of children to show moderate WM capacities (2 score) instead of low WM capacities (1 score; Wald χ^2 (1, N = 118) = 5.26, p < .05; B = 1.41, SE = .62), or a high WM performance (3 score; Wald χ^2 (1, N = 118) = 6.19, p < .05; B = 1.41, SE = .57; *Nagelkerke's* R^2 = .09). The odds of a child with a mother low on postnatal RF obtaining a score of 2 instead of 1 or 3 on the WM task is 4.1 times higher than for a child with a mother A/H on postnatal RF. Again, no such differences were found regarding obtaining a WM score of 1 versus 3 (table 12). Separate multinominal logistic regression indicated that children's withdrawn levels significantly predicted a WM score of 3 versus 2 (Wald χ^2 (1, N = 118) = 5.40, p < .05; B = 2.58, SE = 1.11; *Nagelkerke's* R^2 = .06). The odds ratio shows that when a child's withdrawn level increases by 1 the odds of obtaining a 2 score instead of 3 increases by 13.19. Although the mean withdrawn levels of children with a WM score of 1 and 3 were equal, no significant differences were found regarding withdrawn level and obtaining a 1 versus a 2 score, due to a smaller group size of children with a 1 score (table 13).

Table 12. *WM scores low vs. A/H child-related and postnatal RF*

Measurement	*WM task score*	$N_{child\text{-}related\ RF}$	$N_{postnatal\ RF}$
Low maternal RF	1	2	4
	2	17	17
	3	7	5
	total	26	26
A/H maternal RF	1	30	28
	2	29	29
	3	33	35
	total	92	92

Table 13. *WM score and children's withdrawn level*

Measurement	*WM task score*	*N*	$M_{withdrawn}$	$SD_{withdrawn}$
Child withdrawn level	1	32	.38	.22
	2	46	.49	.18
	3	40	.39	.22

The relationship between early ToM and EF capacities.

First, the relationship between the scores of the three ToM tasks was assessed. Results showed a significant correlation between VP-taking and children's imitation capacities (r = .23, p < .05) and between VP-taking and DD success (r = .22, p < .05). No significant correlation was found between children's imitation capacities and DD performance (r = .07, p = n.s.). Assessing the possible relationship between children's performance of the EF tasks, Cramer's V test of the independence of two categorical variables (one with > 2 categories) showed no significant association between response inhibition and WM performance (Cramer's V (2, N = 118) = .17, p = n.s.). Finally the relations between the ToM tests and the EF tasks were assessed. Passing or failing the delay inhibition task was neither significantly associated with DD performance (χ^2 (1, *N* = 118) = .35, *p* = n.s), nor with imitation capacities (*r* = -.09, *p* = n.s.) or VP-taking abilities (*r* = .03, *p* = n.s.). However, children's scoring on the WM task was significantly and positively related to DD performance (Cramer's V (2, *N* = 118) = .23, *p* < .05). No significant associations were found between WM capacity on the one hand and imitation capacities (Kendall's τ (2, *N* = 118) = .06, *p* = n.s.) and VP-taking on the other hand (Kendall's τ (2, *N* = 118) = .07, *p* = n.s.).

Effectiveness of the early intervention program.

Repeated-measures analysis of variance showed that the mean level of maternal RF increased significantly from T1 (when mothers were in their third trimester of pregnancy and the PI-R was conducted) to T4 (when their children were about 20 months old and the PDI was conducted; (*F* (1, 117) = 21.56, *p* < .001, partial $\eta_p{}^2$ = .16; table 14). Subsequently, we used mixed design analysis of variance to assess whether the increase in maternal RF was different for mothers low on prenatal RF (coaching and non-coaching group) compared to the A/H maternal RF group (coaching and non-coaching group; table 14). As homogeneity of variances of the group distributions could not be assumed, the analysis could not be pursued. However, visual inspection seems to show that the increase in maternal RF does not deviate much across the two intervention groups and between both control groups (figure 4). We proceeded with a comparison of the original experimental groups; the LR-CG, the HR-CG and the HR-IG (table 14). Next to a significant main effect of time (*F* (1, 115) = 22.20, *p* < .001, partial $\eta_p{}^2$ = .16), the results showed a significant main effect of experimental group (*F* (2, 115) = 9.34, *p* < .001, partial $\eta_p{}^2$ = .14). Bonferroni post hoc tests revealed that maternal RF of the HR-CG group and the HR-IG group were significantly lower compared with the LR-CG group. No such differences were found between the HR-CG and HR-IG. Moreover, no significant interaction effect of time and experimental group was present, meaning that no significant differences in increase in maternal RF between the experimental groups were present (F (2, 115) = 1.25, p = n.s.; figure 4). To come up with a possible explanation for the latter finding, the role of maternal age was explored. Mothers' age was significantly correlated with both prenatal RF (r = .26, p < .01) and postnatal RF (r = .20, p < .05), however the increase in maternal RF from T1 to T4 showed not to be significantly related to maternal age (r = -.04, p = n.s.). This indicates that maternal RF increases with age (in the range of 16 to 26), but that the amount of increase is equal for mothers in this age range.

Table 14. *Effectiveness of the early intervention program.*

Measurement	N	M_{RF}	SD_{RF}
Prenatal total RF (T1: PI-R)	118	3.80	.94
Postnatal total RF (T4: PDI)	118	4.25	1.00
Prenatal total RF (T1: PI-R)			
Low coaching	13	2.69	.48
Low, no coaching	35	2.91	.28
A/H coaching	8	4.13	.35
A/H, no coaching	62	4.48	.62
Postnatal total RF (T4: PDI)			
Low coaching	13	3.54	.52
Low, no coaching	35	3.86	.85
A/H coaching	8	4.25	1.04
A/H, no coaching	62	4.61	1.01
Prenatal total RF (T1: PI-R)			
Low risk	60	4.15	.90
High risk control	37	3.54	.84
High risk coaching	21	3.24	.83
Postnatal total RF (T4: PDI)			
Low risk	60	4.45	1.06
High risk control	37	4.16	.93
High risk coaching	21	3.81	.81

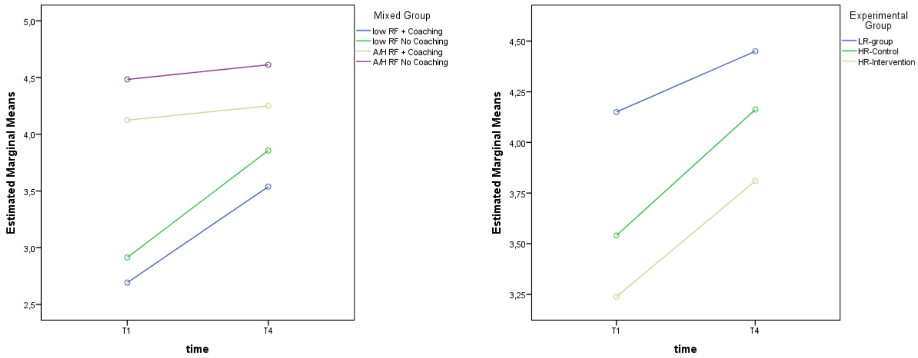

Figure 4. No significant interaction of time (T1 vs. T4) and mixed group (prenatal RF low vs. A/H plus coaching vs. non-coaching) or experimental group (LR, HR-Control and HR-Intervention).

Discussion

Previous studies have found that higher levels of maternal mentalization, measured using MMM, predict better ToM performance at 26, 45, 49 and 51 months of age (Laranjo et al., 2010; Laranjo, Bernier, Meins & Carlson, 2014; Meins et al., 2002; Meins, Fernyhough, Arnott, Leekam & De Rosnay, 2013; Ruffman, Slade, Devitt & Crowe, 2006), that MMM and maternal autonomy support are positively correlated with WM scores at 18 months of age (Bernier et al, 2014), and that a positive parenting style is associated with proper inhibition development (Roskam et al., 2014). The present study aimed to examine the effect of maternal mentalization, operationalized by maternal RF, on ToM and EF capacities of 20-month-old children. In addition, the influence of child gender, temperament and language capacities on these associations and outcomes was assessed. Furthermore, it was examined whether the previously found association between early ToM and EF development of children aged two and older (Carlson & Moses, 2001; Carlson, Mandell & Williams, 2004; Hughes & Ensor, 2007; Hughes & Ensor; Hughes & Ensor, 2007) was also present in our sample of 20 month olds. Last the effectiveness of the early intervention program aimed among others at improving mothers' RF was assessed.

Predicting ToM and EF capacities from maternal RF

VP-taking. As hypothesized, 20 month old children of mothers A/H on child-related or composite RF showed better VP-taking capacities compared with children of mothers low on child-related or composite RF, indicating that mothers who are better able to reflect on their children's behavior, thoughts and feelings and on their relationships with their children, have 20 month olds who are better at simple (level 1) perspective taking (understanding that others need to have their eyes open and directed toward an object, without something blocking their vision, in order to be able to see an object). A trend in the same direction was found using prenatal RF. Although effect sizes are small, these results suggest that mothers' ability to reflect upon their children as separate human beings and upon their relationship with their children, and not mothers' ability to reflect upon their own thoughts, feelings and behaviors, is relevant for the appropriate development of simple VP-taking capacities in their offspring. Furthermore, our findings revealed that, next to the found effect of maternal RF, children's internalizing problems (CBCL) uniquely and significantly predicted VP-taking; children who were reported to have more internalizing problems showed poorer VP-taking performance. In children who are more anxious/depressed, more emotionally reactive and have higher withdrawn levels and somatic complaints, their temperament might interfere with the learning process of acquiring VP-taking skills. Moreover, children with internalizing problems might underperform on the VP task, as the task requires free moving and expression. In line with our expectations and previous research (Rothbart & Ahadi, 1994), we found significant moderating effects of child temperament or problem behavior (attention problems and withdrawn level) regarding the effect of maternal RF on VP-taking, confirming the vulnerability of a difficult temperament for adverse outcomes. Children's high withdrawn levels were associated with poor VP-taking capacities when mothers were low on maternal RF, but not when their mothers had A/H maternal RF capacities. Possibly, mothers low on maternal RF do not show their highly withdrawn offspring the art of perspective taking in a way that suits their child, as they don't sufficiently take into account their child as a separate human being with his/her own temperament. The significant moderating effect of children's attention problems (CBCL) regarding the effect of prenatal RF on VP-taking performance, shows that children with levels of attention problems in the borderline or clinical range need mothers who are able to understand and take into account their difficult temperament (attention problems), in order to adequately apprehend VP-taking. A possible explanation why this moderation effect is only found using prenatal RF, could be that the effects of attention problems on acquiring VP-taking skills start at an early age (as prenatal RF was measured during pregnancy). Alongside the effect of maternal RF, our findings indicate child language capacities as significant positive predictor of VP-taking. As language comprehension and production requirements for the VP task are rather high for 20 month olds, the found effect of language capacities on VP-taking performance can be apprehended. It is expected that these language requirements will be relatively lower at 2.5 years old, when MINDS Leiden will repeat the VP-task. Schaerlaekens' phases of language development indicate that at the end of the early lingual period (from 1 to 2.5 years) language production of normally developing children has developed to using sentences of three to five words (Schaerlaekens, 2008). As language comprehension precedes language production, it can be expected that at 2.5 years most normally developing children fully understand the instructions of the VP-taking task (Maassen & Bol, 2011). Therefore, we expect that future studies of MINDS Leiden, measuring VP-taking at 2.5 years, will reveal a less important role of language capacities and a more profound effect of maternal RF on children's VP-taking capacities.

DD performance. Our findings did not indicate an effect of maternal RF on children's DD understanding. However, language capacities at 20 months significantly predicted DD success. Besides, mothers A/H on maternal RF (prenatal, child-related and composite RF) compared with mothers low on RF seem to have children with significantly better language capacities. A possible explanation of not finding a direct effect of maternal RF on DD performance, is that the DD task was too difficult, causing children with average DD capacities to obtain a zero score, resulting in too little score differentiation. Following this reasoning, it would be interesting to examine whether maternal RF has an effect on DD performance of children 2.5 years old. Future studies of MINDS Leiden should be able to answer this question, as the DD task will be conducted again at T5. Furthermore, the trend regarding the negative effect of children's aggressive behavior (PASEC) on DD success, can be understood as the experimenter three times had to take back the tray with the snacks, and more aggressive children are more likely to protest violently, jeopardizing further instructions.

Imitation capacities. Although imitation capacities of children of mothers A/H on maternal RF were slightly higher than of the low RF group (except for self-related RF), no significant effect of maternal RF on children's imitation capacities was found. A possible explanation is that not enough power was present in the current study, because the low RF group sizes were too small. Moreover, the range of possible scores of the imitation task may have been too limited. Another possibility is that no direct effect of maternal RF on children's imitation abilities is present. In this respect it should be noted that the results indicated that children's imitation performance was positively predicted by child language capacities, a capacity that is influenced by maternal RF group. Furthermore, it was found that children's level of internalizing problems (CBCL) negatively predicted imitation capacities. Children with high internalizing problems (e.g. high withdrawn level) performed worse on the imitation task, possibly because they did not dare to show the necessary gestures and/or noises in the experimental setting, although an effort was made to create a positive atmosphere.

Inhibition. No effect of maternal RF on children's response inhibition capacities was found. Our results showed only one significant predictor of children's response inhibition capacities: child gender. 20-month-old girls proved to be better response inhibitors than boys of this age, a finding that is in line with previous research (Berlin & Bohlin, 2002; Carlson & Moses, 2001; Kochanska et al., 2000; Mileva-Seitz et al., 2015). There are several possible explanations why no effect of maternal RF on response inhibition was found. First, it is possible that parental factors like maintaining discipline instead of maternal mentalization predict response inhibition. This explanation is sustained by research showing that a positive parenting style, higher monitoring and lower discipline, inconsistency and negative controlling are associated with inhibition development (Roskam et al., 2014). Furthermore, no studies could be found specifically linking maternal mentalization to better inhibitory control, neither regarding infants, nor toddlers or older children. Nevertheless, it is possible that our sample consisted of children that were too young to find an effect of maternal RF on response inhibition. As two 'delay of gratification tasks' will be conducted at T5, later results of the MINDS Leiden study may reveal whether an effect of maternal RF on inhibition in older children is present.

WM. Partly confirming and partly contradicting our expectations, we found that mothers low on maternal RF (child-related and postnatal RF) have children who significantly more often show average instead of low or high WM capacities compared to mothers A/H on maternal RF. Although hard to explain, further analyses of possible subgroup differences in temperament or gender, revealed that in our sample children with moderate WM performance are more withdrawn than children with high or low WM capacities. Our previous findings showed that highly withdrawn children perform worse on the VP and imitation task. However, it seems that a higher withdrawn level, to a certain extent, is beneficial for WM performance. Possibly, because low withdrawn levels may lead to a fixation on the objects of the task (e.g. not wanting to return the cups or Ernie), resulting in children paying less attention to the actual instructions and subsequently in a low WM score. It should be noted that in our sample, children with moderate WM capacities and mothers low on maternal RF, seem to be the most withdrawn (although no significant differences were found, probably because group sizes became too small). Another possible, though not satisfactory, explanation is that although significant effects were found at a 95% significance level, the effects were found by chance and do not apply to the population, as the low RF groups were relatively small (N = 26). As MINDS Leiden continues, and more T4 home visits still have to be conducted, a future study may reveal if the effect found holds.

Concluding, we can say that maternal RF, as far as child-related RF aspects are involved, positively seems to predict some aspects of ToM capacities of 20-month-old children. Especially, the effect of maternal RF on children's simple VP-taking capacities proved to be present at this young age. Furthermore, the results confirm that the principle of good enough parenting applies to maternal RF, and in line with the predisposition model, vulnerable children (e.g. children with a difficult temperament or behavior problems) suffer more from low maternal RF abilities. As previous research examining the effect of maternal mentalization on ToM abilities did not use samples consisting of under two year olds, the reason of our non-significant findings regarding both DD and imitation performance could well be our sample of 20 months old children. Future research should examine the proposed effect when children are somewhat older, e.g. at 2.5 years. At this age language capacities are thought to play a less important role and the DD task may become more age appropriate. Following the contradictory effect of maternal RF on children's WM performance and our non-significant findings regarding inhibition, it seems that, at least at this young age, no unequivocal effect of maternal RF on EF capacities is present. Future research should reveal whether maternal RF predicts EF performance at a somewhat older age or whether disciplinary aspects of parenting (Roskam et al., 2014), instead of maternal mentalization, influence EF.

The relationship between early ToM and EF capacities.

The present study revealed only limited support for the association of ToM and EF capacities at 20 months, as a significant (positive) relation was only found between DD and WM performance at this young age. This incomplete finding is in line with previous research as studies show contradicting results regarding the association between ToM and EF in two year olds (Carlson, Mandell & Williams, 2004; Hughes & Ensor, 2005). There are several possible explanations for the results found in the present study. First, our sample may consist of children that are too young to possess certain well established ToM and EF capacities. However, this explanation is contradicted by previous research indicating that simple VP-taking (Luo & Baillargeon, 2007; Poulin-Dubois, Sodian, Metz, Tilden & Schoeppnes, 2007), DD capacities (Repacholi & Gopnik, 1997), imitation abilities (Meltzoff, 1995), inhibitory control (Rothhart, Derryberry & Posner, 1994) and holding information in WM (Diamond & Goldman-Rakic, 1989) are acquired by 18 months. It is also possible that the children in our sample are too young to reliably measure ToM and EF performance. This explanation is sustained by the fact that the downward adaptations of existing tasks, used in the present study, have not been extensively tested. Our findings may also indicate that although 20 month olds possess certain ToM and EF capacities that can be measured properly, these capacities develop independently in the first two years. In this respect it should be noted that no previous research with children under two years of age has indicated a correlation between ToM and EF capacities. Finally, an explanation could be that in the present study delay inhibition was examined given children's young age, although research reveals that more profound associations are found using conflict or cognitive inhibition (Carlson, Claxton & Moses, 2015; Carlson & Moses, 2001; Dennis Agostino, Roncadin & Levin, 2009). There are several explanations regarding the observed significant association between children's DD and WM performance. Possibly, children need WM capacity to keep in mind what snack the experimenter prefers, so they can offer the right snack. However, as the majority of literature indicates a functional interdependence (Carlson et al., 2015; Dennis et al., 2009; Hughes & Ensor, 2007), it seems more likely that higher levels of acquired executive skills enable a better understanding of DD. Another possible explanation could be that the scores of both tasks is influenced by the extent to which children are able to go with the flow of the tasks and let the experimenter take back the objects and snacks after each trial, before the next trial can take place. This ability to go with the flow might be influenced by both children's withdrawn level, that predicted DD success, and by aggressive behavior (PASEC), that seemed to affect WM performance. During the VP- and imitation task, if necessary, children were offered the next toy before they had to return the previous one. This distracting procedure was not possible with respect to the DD and WM task.

Effectiveness of the early intervention program.

No support was found for our hypothesis that maternal RF would increase more from T1 to T4 for mothers in the intervention group, compared to mothers in the control group(s). Instead, the results showed that maternal RF significantly increased over this time period for the HR-IG as well as for the HR-CG and the LR-CG. Prior research examining the effect of a similar intervention like the one used in the current study, called Minding the Baby, also could not confirm the effectiveness of the intervention with respect to increasing maternal RF (Ordway et al., 2014). There are several possible explanations of the present study revealing that maternal RF increased significantly prenatally to postnatally, however, not finding a significant effect of the coaching program on maternal RF. Maybe there is a profound 'natural' increase in maternal RF from T1 to T4 of the relatively young, primiparous women of the current sample, making it more difficult for an intervention to increase RF even further in this time frame. Maternal age is a natural cause of the increase in maternal RF, as we found significant positive correlations between maternal age and RF. Although not investigated, it seems likely that this correlation is less profound regarding mothers aged 25 or above. Besides, becoming a mother for the first time (and becoming responsible for another human being) might also naturally activate RF. This activating mechanism might be even more at work regarding relatively young women, who before did not assume responsibility for their lives. Another explanation of our disappointing results could be that although the intervention, among others, aimed at improving maternal RF, in practice most time and effort have been spent on improving mothers' life style, as acute problems might have arisen frequently in HR-IG during the 2.5 years the intervention took place. In addition, the random assignment procedure regarding the HR-IG and the HR-CG could not be consequently followed, making it a possible cause for bias.

Strengths and limitations

The present study contributes to the insight regarding the effect of maternal RF on precursors of ToM and EF and with respect to the association between ToM and EF in 20-month-old children. The fact that the sample of the present study consisted of younger children compared to previous research examining the role of maternal mentalization on ToM and EF, and assessing the association between ToM and EF development, results in a unique contribution of the current study. Nevertheless, the young age of the children also constitutes a limitation of our research, as it remains a challenge to measure precursors of Tom and EF in a valid and reliable manner. A follow-up study should examine the effect of maternal RF on ToM and EF capacities in somewhat older children (e.g. 2.5 years). As MINDS Leiden is a prospective, longitudinal study from pregnancy onwards, prenatal- as well as postnatal measures of maternal RF were incorporated in the present study. Furthermore, the study used sub measures of postnatal RF (child-related and self-related RF) as well as a composite RF score, making it possible to determine what specific type of maternal RF affected ToM and EF abilities.

The relatively small size of the subgroups created regarding maternal RF (low vs. A/H) and child temperament (low vs. BC), and of the HR-IG, is considered a weakness of the present study. Moreover, the random assignment of high risk mothers to either the HR-IG or the HR-CG could not always be acted upon, as some high risk mothers refused to accept the coaching program and others had to be offered the coaching program in order for them to participate in the study. Another limitation of the present study is that child temperament was examined using questionnaires, which were only filled in by the mother. Our results, indicating that mothers high on self-related maternal RF were more likely to report their children to have a difficult temperament, point to a possible reporting bias.

Future directions and implications

Our finding that mothers with low maternal RF capacities have children who show less simple VP-taking abilities at 20 months, and that mothers at risk have lower maternal RF capacities, confirms the importance of early intervention programs to focus on improving maternal RF of these mothers at risk. Following the result that especially child-related aspects of RF affected this aspect of early ToM, intervention programs should concentrate on mothers' representations of their children and of their relationships with their children, and less on mothers' reflections upon their own thoughts, feelings and behavior. Furthermore, maternal RF plays an even more important role when children with a difficult temperament are at stake. This implies that therapists of coaching programs should especially pay attention to improving maternal RF when coming across these vulnerable mother-child dyads. Future research should reexamine the effect of maternal RF on ToM and EF capacities, and the relationship between ToM and EF, in somewhat older children, e.g. 2.5 years, as the current study did not find the proposed effects on EF, DD and imitation abilities, possibly due to difficulties measuring ToM and EF in 20 month olds, and/or because

ToM and EF capacities might not have been established well enough at 20 months. In this respect, it is recommended that future research will examine possible main- and moderating effects of child temperament, gender and language, given the results of the current study. Although the present study showed disappointing findings regarding the effectiveness of the early intervention program on improving maternal RF, future studies (of MINDS Leiden) should reexamine the effectiveness using a larger intervention group. Besides, the other proposed effects of the early intervention program (improving mothers' life style and social network) should be assessed. The results of these future studies, plus information from the therapists about the amount of time spent on each of the goals of the intervention, can contribute to our understanding in what way the current early intervention program should be adapted to become (more) effective.

References

Achenbach, T.M. & Rescorla, L.A. (2000). *Manual for the ASEBA Preschool forms and Profiles.* Burlington, VT, University of Vermont Department of Psychiatry.

Alink, L.R., Mesman, J., Van Zeijl, J., Stolk, M.N., Juffer, F., Koot, H.M., … Van IJzendoorn, M.H. (2006). The Early Childhood Aggression Curve: Development of Physical Aggression in 10- to 50-Month-Old Children. *Child Development, 77*, 954-966.

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

Anzman-Frasca, S., Stifter, C.A., Paul, I.M. & Birch, L.L. (2014). Negative Temperament as a Moderator of Intervention Effects in Infancy: Testing a Differential Susceptibility Model. *Prevention Science, 15,* 643-653. Doi: 10.1007/s11121-013-0408-4.

Barkley, R.A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin, 121*, 65-94.

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

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (1993). *Understanding other minds: Perspectives from autism*. Oxford: Oxford University Press.

Berlin, L. & Bohlin, G. (2002). Response Inhibition, Hyperactivity, and Conduct Problems Among Preschool Children. *Journal of Clinical Child Psychology, 31*, 242-251.

Bernier, A., Carlson, S.M. & Whipple, N. (2010). From external regulation to self-regulation: Early parenting precursors of young children's executive functioning. *Child Development, 81,* 326-339. Doi: 10.1111/j.1467-8624.2009.01397.x.

Blair, K.A., Denham, S.A., Kochanoff, A. & Whipple, B. (2004). Playing it cool: Temperament, emotion regulation, and social behavior in preschoolers. *Journal of School Psychology, 42*, 419-443. Doi: 10.1016.j.jsp.2004.10.002.

Blair, C., Raver, C.C. & Berry, D.J. (2014). Two Approaches to Estimating the Effect of Parenting on the Development of Executive Function in Early Childhood. *Developmental Psychology, 50*, 554-565. Doi: 10.1037/a0033647.

Bull, R., Phillips, L.H, & Conway, C.A., (2008). The role of control functions in mentalizing: Dual-task studies of Theory of Mind and executive function. Cognition, 107, 663-672. Doi: 10.1016/j.cognition.2007.07.015.

Calero, C.I., Salles, A., Semelman, M. & Sigman, M. (2013). Age and gender dependent development of Theory of Mind in 6- to 8-years old children. *Frontiers in Human Neuroscience, 7*, 281. Doi: 10.3389/fnhum.2013.00281.

Carlson, S.M., Claxton, L.J., & Moses, L.J. (2015). The Relation Between Executive Function and Theory of Mind is More Than Skin Deep. *Journal of Cognition and Development, 16,* 186-197. Doi: 10.1080/15248372/2013.824883.

Carlson, S.M., Mandell, D.J. & Williams, L. (2004). Executive Function and Theory of Mind: Stability and Prediction From Ages 2 to 3. *Developmental Psychology, 40*, 1105-1122. Doi: 10.1037/0012-1649.40.6.1105.

Carlson, S.M. & Moses, L.J. (2001). Individual Differences in Inhibitory Control and Children's Theory of Mind. *Child Development, 72*, 1032-1053.

Davidson, M.D., Amso, D., Anderson, L.C. & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition and task switching. *Neuropsychologia, 44,* 2037-2078.

Dennis, M., Agostino, A., Roncadin, C. & Levin, H. (2009). Theory of mind depends on domain-general executive functions of working memory and cognitive inhibition in children with traumatic brain injury. Journal of Clinical and Experimental Neuropsychology, 31, 835-847. Doi: 10.1080/13803390802572419.

Diamond, A. (2013). Executive Functions. *Annual Review of Psychology, 64,* 135-168.

Diamond, A. & Goldman-Rakic, P.S. (1989). Comparison of human infants and rhesus monkeys on Piaget's AB task: Evidence for dependence on dorsolateral prefrontal cortex. *Experimental Brain Research, 74,* 24-40.

Eigsti, I., Zayas, V., Mischel, W., Shoda, Y., Ayduk, O., Dadlani, M.B., Casey, B.J. (2006). Predicting cognitive control from preschool to late adolescence and young adulthood. *Psychology Science, 17*, 478-484.

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

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

Fonagy, P., Steele, H., Moran, G, Steele, M. & 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, 13*, 200-217.

Fonagy, P. & Target, M. (2005). Bridging the transmission gap: And end to an important mystery of attachment research? *Attachment and Human Development, 7*, 333-343.

Frith, C.D. & Frith, U. (2006). The neural basis of mentalizing. *Neuron 50,* 531-534. Doi 10.1016/j.neuron.2006.05.001.

Frith, U. & Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society B: Biological Sciences, 358,* 459-473. Doi 10.1098/rstb.2002.1218.

Garstein, M.A. & Rothbart, M.K. (2003). Studying infant temperament via the revised infant behavior questionnaire. *Infant Behavior & Development, 26,* 64-86. Doi: 10.1016/S0163-6383(02)00169-8.

Geurts, H.M. & Huizinga, M. (2011). Aandacht en executieve functies. In H. Swaab, A. Bouma, J. Hendriksen & C. König (Eds.), *Klinische kinderneuropsychologie.* (pp. 169-188). Amsterdam: Boom.

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 behavior in the intergenerational transmission of attachment. *Attachment & Human Development, 7,* 299-311.

Griffin, K.W., Botvin, G.J., Scheier, L.M., Diaz, T. & Miller, N.L. (2000). Parenting Practices as Predictors of Substance Use, Delinquency, and Aggression Among Urban Minority Youth: Moderating Effects of Family Structure and Gender. *Psychology of Addictive Behaviors, 14*, 174-184.

Hammond, S.I., Müller, U., Carpendale, J.I.M., Bibok, M.B. & Liebermann-Finestone, D.P. (2012). The Effects of Parental Scaffolding on Preschoolers' Executive Function. *Developmental Psychology, 48*, 271-281. Doi: 10.1037/a0025519.

Happé, F., Booth, R., Charlton, R. & Hughes, C. (2006). Executive function deficits in autism spectrum disorders and attention-deficit/hyperactivity disorder: Examining profiles across domains and ages. *Brain and Cognition, 61,* 25-39. Doi: 10/1016/j.bandc.2006.03.004.

Happé, F. & Frith, U. (1996). Theory of mind and social impairment in children with conduct disorder. *British Journal of Developmental Psychology, 14,* 385-398.

Harris, P.L. (1992). From Simulation to Folk Psychology: The Case for Development. Mind & Language, 7, 120-144.

Hirshfeld-Becker, D.R., Biederman, J., Calltharp, S., Rosenbaum, E.D., Faraone, S.V. & Rosenbaum, J.F. (2003). Behavioral Inhibition and Disinhibition as Hypothesized Precursors to Psychopathology: Implications for Pediatric Bipolar Disorder. *Biological Psychiatry, 53*, 985-999. Doi: 10.1016.S0006-3223(03)00316-0.

Horton, M.K., Kahn, L.G., Perera, F., Barr, D.B. & Rauh, V. (2012). Does the home environment and the sex of the child modify the adverse prenatal exposure to chlorpyrifos on child working memory? *Neurotoxicology and Teratology, 34*, 534-541. Doi: 10.1016/j.ntt.2012.07.004.

Hughes, C., Jaffee, S.R., Happé, F., Taylor, A., Caspi, A. & Moffit T.E. (2005). Origins of Individual Differences in Theory of Mind: From Nature to Nurture?

Hughes, C. & Ensor, R. (2005). Executive Function and Theory of Mind in 2 Year Olds: A Familiy Affair? *Developmental Neuropsychology, 28,* 645-668. Doi: 10.1207/s15326942dn2802_5.

Hughes, C. & Ensor, R. (2007). Executive Function and Theory of Mind: Predictive Relations From Ages 2 to 4. *Developmental Psychology, 43,* 1447-1459. Doi: 10.1037/0012-1649.43.6.1447.

Ivanova, M.Y., Achenbach, T.M, Rescorla, L.W., Harder, V.S., Ang, R.P., Bilenberg, N.,… Verhulst, F.C. (2010). Preschool Psychopathology Reported by Parents in 23 Societies: Testing the Seven-Syndrome Model of the Child Behavior Chicklist for Ages 1.5-5. *Journal of the American Academy of Child & Adolescent Psychiatry, 49,* 1215- 1224.

Jessee, A., Mangelsdorf, S.C. & Wong, M.S. (2012). Temperament as a Moderator of the Effects of Parental Depressive Symptoms on Child Behavior Problems. *Social Development, 21*, 610-627. Doi: 10.1111/j.1467-9507.2011.00639.x.

Kochanska, G., Murray, K., Jacques, T.Y., Koenig, A.L. & Van de Geest, K.A. (1996). Inhibitory Control in Young Children and Its Role in Emerging Internalization. *Child Development, 67*, 490-507.

Kochanska, G., Murray, K.T. & Harlan, E.T. (2000). Effortful Control in Early Childhood: Continuity and Change, Antecedents, and Implications for Social Development. *Developmental Psychology, 36*, 220-232. Doi: 10.1037/0012-1649.36.2.220.

Kok, R., Lucassen, N., Bakermans-Kranenburg, M.J., Van IJzendoorn, M.H., Ghassabian, A., Roza, S.J., …Tiemeier, H. (2014). Parenting, corpus callosum, and executive function in preschool children. *Child Neuropsychology, 20,* 583-606. Doi: 10.1080/09297049.2013.832741.

Laranjo, J., Bernier, A., Meins, E. & Carlson, S.M. (2010). Early Manifestations of Children's Theory of Mind: The Roles of Maternal Mind-Mindedness and Infant Security of Attachment. *Infancy, 15*, 300-323. Doi: 10.1111/j.1532-7078.2009.00014.x.

Laranjo, J., Bernier, A., Meins, E. & Carlson, S.M. (2010). The roles of maternal mind-mindedness and infant security of attachment in predicting preschoolers' understanding of visual perspective taking and false belief. *Journal of Experimental Child Psychology, 125*, 48-62. Doi: 10.1016/j.jecp.2014.02.005.

Leslie, A.M. (1987). Pretense and representation: the origins of "theory of mind". *Psychological Review, 94,* 412-426.

Lord, C., Risi, S., Lambrecht, L., Cook Jr., E.H., Leventhal, B.L., DiLavore, P.C., Pickles, A. & Rutter, M. (2000). The Autism Diagnostic Observation Schedule-Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *Journal of Autism and Developmental Disorders, 30*, 205-223.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L & Schopler, E. (1989). Autism Diagnostic Observation Schedule: A Standardized Observation of Communicative and Social Behavior. *Journal of Autism and Developmental Disorders, 19*, 185-212.

Luo, Y. & Baillargeon, R. (2007). Do 12.5-month-old infants consider what objects others can see when interpreting their actions? *Cognition, 105*, 489-512. Doi: 10.1016.j.cognition.2006.10.007.

Maassen, B. & Bol, G. (2011). Taal en Spraak. In H. Swaab, A. Bouma, J. Hendriksen & C. König (Eds.), *Klinische kinderneuropsychologie.* (pp. 123-146). Amsterdam: Boom.

McFadyen-Ketchum, S.W., Bates, J.E., Dodge, K.A. & Pettit, G.S. (1996). Patterns of Change in Early Childhood Aggressive-Disruptive Behavior: Gender Differences in Predictions from Early Coercive and Affectionate Mother-Child Interactions. *Child Development, 67*, 2417-2433.

Meins, E., Fernyhough, C., Wainwright, R., Das Gupta, M., Fradley, E. & Tuckey, M. (2002). Maternal mind-mindedness and attachment security as predictors of theory of mind understanding. *Child Development, 73*, 1715-1726.

Meins, E., Fernyhough, C., Arnott, B., Leekam, S.R. & De Rosnay, M. (2013). Mind-Mindedness and Theory of Mind: Mediating Roles of Language and Perspectival Symbolic Play. Child Development, 84, 1777-1790. Doi: 10/1111/ cdev.12061.

Mileva-Seitz, V.R., Ghassabian, A., Bakermans-Kranenburg, M.J., Van den Brink, J.D., Linting, M., Jaddoe, V.W.V., … Van IJzendoorn, M.H. (2015). Are boys more sensitive to sensitivity? Parenting and executive function in preschoolers. *Journal of Experimental Child Psychology, 130,* 193-208. Doi: 10.1016/j.jecp.2014.08.008.

Meltzoff, A.N. (1995). Understanding the intentions of others re-enactment of intended acts by 18 month-old children. Developmental Psychology, 31, 838-850.

Moffitt, T.E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R.J., Harrington, H. & Caspi, A. (2011). A gradient of childhood self-control predicts health, wealth, and public safety. *Proc. Natl. Acad. Sci. USA 108,* 2693-2698.

Moll, H. & Meltzhoff, A.N. (2011). How Does It Look? Perspective-Taking at 36 Months of Age. *Child Development, 82*, 661-673.

Ordway, M.R., Sadler, L.S., Faan, P.C., Dixon, J., Close, N., Mayes, M.D. & Slade, A. (2014). Lasting effects of an interdisciplinary home visiting program of child behavior: preliminary follow-up results of a randomized trial. *Jounal of Pediatric Nursing, 29,* 3-13. Doi: 10.1016/j.pedn.2013.04.006.

Ozonoff, S., Pennington, B. & Rogers, S. (1991). Executive function deficits in high-functioning autistic children: Relationship to theory of mind. *Journal of Child Psychology and Psychiatry, 31,* 1081-1105.

Perner, J. & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences, 3,* 337- 344.

Perner, J., Lang, B., & Kloo, D. (2002). Theory of Mind and Self-Control: More than a Common Problem of Inhibition. *Child Development, 73,* 752-767.

Poulin-Dubois, D., Sodian, B., Metz, U., Tilden, J. & Schoeppner, B. (2007). Out of Sight Is Not Out of Mind: Developmental Changes in Infants' Understanding of Visual Perception During the Second Year. *Journal of Cognition and Development, 8*, 401-425.

Putnam, S.P., Gartstein, M.A. & Rothbart, M.K. (2006). Measurement of fine-grained aspects of toddler temperament: The Early Childhood Behavior Questionnaire. *Infant Behavior & Development, 29*, 386-401. Doi: 10.1016.j.infbeh.2006.01.004.

Repacholi, B.M. & Gopnik, A. (1997). Early Reasoning About Desires: Evidence From 14- and 18-Month-Olds. *Developmental Psychology, 33*, 12-21.

Roskam, I., Stievenart, M., Meunier, J., & Noël, M. (2014). The development of children's inhibition: Does parenting matter? *Journal of Experimental Child Psychology, 122,* 166-182. Doi 10.1016/j.jecp.2014.01.003.

Rothbart, M.K., Derryberry, D. & Posner, M.I. (1994). *A psychobiological approach to the development of temperament.* In J.E. Bates & T.D. Wachs (Eds.), Temperament: Individual differences at the interface of biology and behavior (pp. 83–116). Washington, DC: American Psychological Association.

Rothbart, M.K. & Ahadi, S.A. (1994). Temperament and the Development of Personality. *Journal of Abnormal Psychology, 103*, 55-66.

Rothbart, M.K. & Jones, L.B. (1998). Temperament, self-regulation, and education. *School Psychology Review, 27*, 479-492.

Rothbart, M.K., Ahadi, S.A., Hershey, K.L., & Fisher, P. (2001). Investigations of Temperament at Three to Seven Years: The Children's Behavior Questionnaire. Child Development, 72, 1394-1408.

Ruffman, T., Slade, L., Devitt, K. & Crowe, E. (2006). What mothers say and what they do: The relation between parenting, theory of mind, language and conflict/cooperation. *British Journal of Developmental Psychology, 24*, 105-124. Doi: 10.1348/026151005X82848.

Schaerlaekens, A. (2008). *De taalontwikkeling van het Kind.* Groningen/Houten: Wolters-Noordhoff.

Schaerlaekens, A. (1995). Einschätzen und Messen der Sprachfähigkeit beim Kinder. *Folia Phoniatr. Logop., 47*, 303-309.

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,* 737- 754. Doi: 10/1111/j.1467-9507.2007.00457.x

Sharp, C., Fonagy, P. & Goodyer, I. (2006). Imagining your child's mind: Psychosocial adjustment and mothers' ability to predict their children's attributional response styles. *British Journal of Developmental Psychology, 24*, 197-214.

Sharp, C., Fonagy, P. & Goodyer, I. (2008). *Social Cognition and Developmental Psychopathology.* Oxford: Oxford University Press.

Slade, A. (2005). Parental reflective functioning: an introduction. *Attachment & human development, 7*(3), 269-281.

Slade, A. (2010). Reflective Parenting Programs: Theory and Development. Psychoanalytic Inquiry, 26, 640-657. Doi: 10/1080/07351690701310698.

Slade, A., Bernbach, E., Grienenberger, J., Levy, D., & Locker, A. (2005). *Addendum to reflective functioning scoring manual (Fonagy, Steele, Steele, & Targer, 1998) for use with the Parent Development Interview*. New York: The City College of NewYork.

Slade, A., Grienenberger, J., Bernbach, E., Levy, D. & Locker, A. (2005). Maternal reflective functioning, attachment, and the transmission gap: A preliminary study. *Attachment & Human Development, 7*(3), 283-298. doi:10.1080/14616730500245880

Slade, A., Patterson, M., & Miller, M. (2007). *Addendum to Reflective Functioning Scoring Manual (Fonagy, Steele, Steele, & Target, 1998) for use with the Pregnancy Interview*. New York: City College of New York.

Smaling, H. (2013). *Parent Developmental Interview (PDI)- herziene versie (Nederlandse vertaling van het Parent Development Interview – revised)*. Universiteit Leiden. Leiden.

Smaling, H. & Suurland, J. (2011). *Zwangerschapsinterview - herziene versie (Nederlandse vertaling van het Pregnancy Interview - revised)*. Universiteit Leiden. Leiden.

Smaling, H.J.A., Huijbregts, S.C.J., Suurland, J., Van der Heijden, K.B., Van Goozen, S.H.M. & Swaab, H. (2015). Prenatal reflective functioning in primiparous women with a high-risk profile. Infant Mental Health Journal, 1, 1-11. Doi: 10.1002/imhj.21506.

Spencer, T.J. (2006). ADHD and comorbidity in childhood. *Journal of Clinical Psychiatry, 67,* 27-31.

Swaab, H., Bouma, A., Hendriksen, J. & König, C. (2011). *Klinische kinderneuropsychologie.* Amsterdam: Boom.

Walker, S. (2005). Gender Differences in the Relationship Between Young Children's Peer-Related Social Competence and Individual Differences in Theory of Mind. *The Journal of Genetic Psychology, 166*, 297-312.