Paper I

Landsem IP, Handegård BH, Ulvund SE, Tunby J, Kaaresen PI, Rønning JA. Does an early intervention influence behavioral development until age nine in children born prematurely? *Child Development* 2015; 86(4):1063-1079.

CHILD DEVELOPMENT

Child Development, July/August 2015, Volume 86, Number 4, Pages 1063–1079

Does An Early Intervention Influence Behavioral Development Until Age 9 in Children Born Prematurely?

Inger Pauline Landsem
University Hospital of North Norway and UiT, The Arctic University of Norway

Bjørn Helge Handegård
UiT, The Arctic University of Norway

Stein Erik Ulvund
UiT, The Arctic University of Norway and University of Oslo

Jorunn Tunby
University Hospital of North Norway

Per Ivar Kaaresen, and John A. Rønning
University Hospital of North Norway and UiT, The Arctic University of Norway

This study examined whether the Mother–Infant Transaction Program prevents behavioral problems among preterm children (birth weight < 2000 g) until age 9. The program was administered to 72 preterms, while 74 preterms and 75 full-terms formed control groups (N = 221). Behavior was reported by parents (Child Behavior Checklist) and teachers (Teachers Report Form) and by all on selected Strengths and Difficulties Questionnaire (SDQ) questions. Long-term behavioral development appeared to be qualitatively unaffected by the intervention. At ages 7 and 9, fewer attention problems and better adaptation to school were reported from parents and teachers of the intervention group compared to preterm controls. At age 9, teachers reported fewer difficulties in the intervention group and better academic performance. In these areas they were reported as being at the statistically same level as term controls.

Children born preterm are at increased risk of developmental problems (Shonkoff & Phillips, 2000). Prematurity is correlated with behavioral difficulties and a higher incidence of problems in everyday life that persists throughout primary school (Bhutta, Cleves, Casey, Cradock, & Anand, 2002). A higher incidence of attentional deficits is found to be the most obvious problem, but internalizing problems as reported by parents and externalizing behavior as reported by teachers also occur (Aarnoudse-Moens, Weisglas-Kuperus, van Goudoever, & Oosterlaan, 2009). Problems are related to deficits in several cognitive functions and increases in learning problems, especially in disciplines such as mathematics, reading, and spelling (Aarnoudse-Moens et al., 2009). Problems are discovered to be inversely related to birth weight (BW) and probably more frequent among male born children. Between 50% and 60% of premature born children are reported to require special learning assistance (Aylward, 2005; Bhutta et al., 2002). The prevalence of behavioral problems at 8 years is reported to be about 20%, which is twice the reported rate in the general population (Bongers, Koot, van der Ende, & Verhulst, 2003; Gray, Indurkhya, & McCormick, 2004). A Norwegian study found that 40% of preterm born children displayed abnormal behavior at age 11 (Elgen, Holsten, & Odberg, 2012; Elgen, Sommerfelt, & Markestad, 2002). They were also 3 times as likely to receive a diagnosis of psychiatric disorder as their term peers, in accordance with other studies (Johnson, 2007).

Several interventions have been tested in attempts to prevent the development of problem behavior (Glazebrook et al., 2007; McAnulty et al.,

This study was funded by the Northern Regional Health Authority in Norway. We thank children and parents for their participation throughout childhood and Professor Lauritz B. Dahl for his important contributions to the study until his retirement. We are also grateful to the neonatal care nurses, Unni Byrknes, Nina B. Cheetham, Randi Berg Høier, Hilde Sletvold, Anne Uhre, and Kirsti Vikhammer, who together with Tunby and Landsem carried out the intervention program.

Correspondence concerning this article should be addressed to Inger Pauline Landsem, Child & Adolescent Department, Postbox 43, University Hospital of Northern Norway, 9038 Tromsø, Norway. Electronic mail may be sent to inger.pauline.landsem@unn.no.

 0009-3920/2015/8604-0006
DOI: 10.1111/cdev.12368

2010; McCarton et al., 1997; Meijssen et al., 2010; Newnham, Milgrom, & Skouteris, 2009), documenting some immediate effects. However, long-term follow-ups are few (Guralnick, 2012). Only one small study has reported successful intervention effects on preterms that persisted until 9 years of age (Achenbach, Howell, Aoki, & Rauh, 1993) and the current study is a modified replication of their design.

The theoretical foundation of these studies is a transactional understanding of child development (Sameroff & Fiese, 2000). The developing child is an active participant in the creation of his or her interhuman environments. This is formed by the contextual interplay between individual biological conditions (child) and the child's interactions with the environment (family, school). As stated by Hauser-Cram et al. (2001), "Development occurs within multiple hierarchical contexts that are reciprocally related" (p. 21). When children are born prematurely, the natural transactions are easily interrupted because of these infants' immature expressions and limited capacity to self-regulate and emit recognizable alertness responses (Feldman, 2006). This limits the possibilities of caregivers to establish development-supportive caretaking (Nugent, Keefer, Minear, Johnson, & Blanchard, 2007).

The emotional crisis affecting many parents of preterms is also known to influence development (Abidin, 1995; Nurcombe et al., 1984). This study included an initial session when parents could vent feelings such as grief, anger, or frustration related to the preterm delivery, the hospital stay, and how this condition had affected their lives. The reinforcement of parental self-esteem and confidence was another aim of the program as parental experiences and discoveries of each child's developmental expression were highly acknowledged throughout the sessions. The parental guidance was given in steps as the systems of newborn behavior (physiological, motor, state, and emotional regulation) were explored one at the time and progressively put together.

Parents of preterms may also find it difficult to emotionally attach to their newborn. Interactions between parent and child may suffer because of lack of feedback from the child and experiential avoidance is described as a significant predictor of weakened maternal involvement and responsiveness (Evans, Whittingham, & Boyd, 2012). This was taken into account already in the first session, which was dedicated to enhancing the parent–child relationship by waiting for a moment where the child was able to enter a quite alert state and supported to display the best alertness possible. In the following sessions, parents were helped to understand how early biological regulation continuously interacts with the child's availability and self-regulation. Homeostatic stability, sleep–waking cyclic patterns, and the maturity of state regulation are closely associated with the infant's later patterns of interaction (Feldman, 2006). The child's regulatory competence was explored by helping parents recognize signs of well-organized systems versus arousal and signs of stress. At the same time they were helped to make adequate adjustments to these signs and to discover changes as the child matured.

Several mechanisms are likely to be involved in behavior regulation throughout childhood. In infancy, a more flexible maturation from biological rhythms to early social rhythms strengthens the parent–child synchrony (Feldman, 2006) and makes more finely tuned face-to-face interactions possible. This has been shown to strengthen the child's self-regulatory competence (Feldman, Greenbaum, & Yirmiya, 1999; Olafsen et al., 2006; Olafsen et al., 2012). Improved self-regulation is a predictor of fewer behavioral problems in preschool children (Feldman, 2009), and the beneficial effect was assumed to continue into middle childhood years, as has been documented for children diagnosed with disabilities (Hauser-Cram et al., 2001). This study hypothesized that the early intervention would improve long-term behavioral and socio-emotional development in premature children. This hypothesis was supported by intervention effects already published: less parenting stress reported until age 2, improved regulatory competence in infancy and significantly improved cognitive, and behavioral outcomes at age 5 (Kaaresen, Rønning, Ulvund, & Dahl, 2006; Nordhov et al., 2010; Nordhov, Rønning, Ulvund, Dahl, & Kaaresen, 2012; Olafsen et al., 2008; Olafsen et al., 2012). In the Vermont study, Rauh, Achenbach, Nurcombe, Howell, and Teti (1988) suggested that the intervention had fostered transactional patterns that began to interact with the children's development after infancy, characterized as a possible "sleeper effect" in a later study (Achenbach et al., 1993). We lean upon later research showing how the development of regulative behavior seems to be crucial in overall child development (Feldman, 2006, 2009).

This study investigated the long-term intervention effects on behavioral development of preterm children until age 9. Both longitudinal and cross-sectional investigations were incorporated in the

analysis. The following questions were asked: (a) Does an early intervention influence longitudinal behavioral-emotional development, as reported by parents and teachers until age 9? (b) Does an early intervention influence behavioral outcomes, social competence, and adaptive behavior in school as perceived by parents and teachers at ages 7 and 9? (c) Are specific BW groups or genders particularly affected by the intervention? (d) How does the behavioral development of preterms (with and without intervention) compare to the behavioral development of full-term children?

Method

Participants

The Tromsø Intervention Study on Preterms (TISP) is a randomized, controlled study of preterm infants with BW < 2000 g, born at the University Hospital of North Norway between March 1999 and September 2002 (Rønning, Dahl, Ulvund, Kaaresen, 1999). Decisions on sample size were based on the results of a previous study by Achenbach et al. (1993) and were calculated to detect a group difference in Bayley's Mental Developmental Index (MDI; Bayley, 1993) at age 2 of 7.5 points (corresponding to a 0.5 *SD* difference) with 80% probability (Kaaresen et al., 2006). Preterm infants were randomized in blocks of six to form a preterm intervention group (PI; $n = 72$) and a preterm control group (PC; $n = 74$), stratified within gestational age (GA) < 28 weeks. Infants of non-Norwegian-speaking parents and infants diagnosed with severe neurological or sensory impairments were excluded (Figure 1). Children identified with disabilities at later follow-ups continued in the study if able to participate in the age-appropriate assessments. Subanalyses, with children nearly reaching this criterion, were included and will be reported. Healthy newborn infants (GA ≥ 37 weeks) were recruited from the maternity ward at the same hospital as a term reference group (TR; $n = 75$). The PC group followed the Neonatal Intensive Care Unit (NICU) guidelines for discharge of preterm infants. The TR group was examined once by a pediatrician, routinely on the 3rd day of life. Baseline data for each group (Table 1) have previously been described in detail (Kaaresen et al., 2006).

Intervention

The intervention program was a modified version of the Mother–Infant Transaction Program (MITP; Rauh, Nurcombe, Achenbach, & Howell, 1990). Each intervention consisted of seven 1-hr sessions with parents and their baby during the last week before discharge, and four home visits 1, 2, 4, and 12 weeks postdischarge. All sessions were performed in a separate, quiet room with the baby, mother, father (when available), and the intervention nurse. One exception was the initial session where parents met the nurse without the baby present. The interventions were adapted to each family's needs, but strictly followed the topics as described in the MITP manual (Rauh et al., 1990). The seven sessions in the hospital had different agendas, which can briefly be listed as: (a) an exploration of the infant's regulation and social competencies; (b) signs of stress and stability in the homeostatic system and (c) the motoric system, with focus on tone, posture, and movements; (d) the infant's regulation of states and transition's between these; (e) how parents could help the infant become alert and available for interactions; and (f) how to use this knowledge in daily caretaking such as bathing. In the last hospital session all topics were discussed, the parents asked questions, and elements were repeated if asked for.

The first home visit dealt with the families' adaptation to the new environment and how they made adjustments with the infants' needs in mind. In the next visit, parent–child social interactions were addressed and parents' stories about new activities were applauded. The third visit, 1 month after discharge, focused on parents' observations of their infant's behavioral style and on how they could adapt to the rhythms and capacities of their infant. In the last visit at 3 months postdischarge, the program was summed up and changes that had appeared were reviewed with the parents. At this time most of the families expressed that they had identified some useful rhythms and routines in their daily life.

Finalizing the program at this age was justified by the ending of the original study and because it made up an intervention that could realistically be implemented with the financial and professional resources available. Mothers participated in all sessions, while fathers participated on average in 6 of the 12 sessions (interquartile range = 4–10 sessions). Eight nurses were trained to perform the interventions, and each family was followed by the same nurse throughout all the sessions with no additional contact afterward. Unlike the original study, no logbooks were given to the families after the interventions were finished. Completion of the intervention

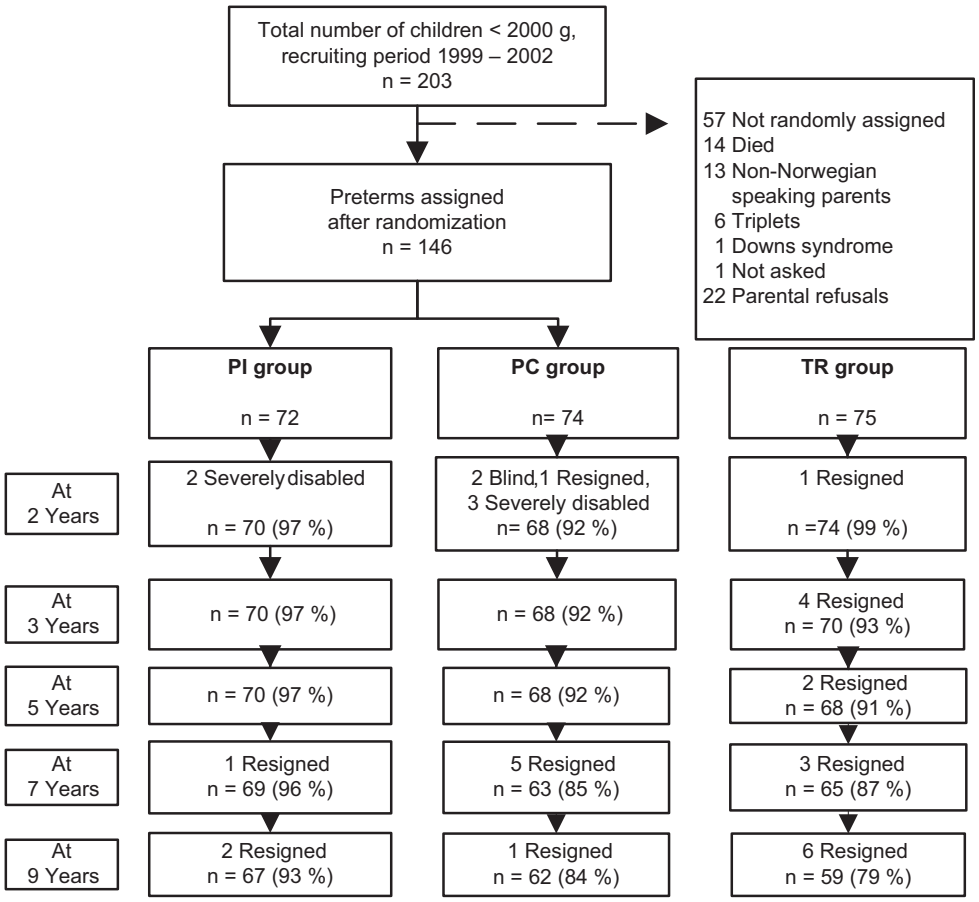

Figure 1. Study flow diagram from ages 2 to 9. PI = preterm intervention; PC = preterm control; TR = term reference.

was ensured by review of logbooks by the study director (J. A. Rønning).

Instruments

Mothers and fathers reported independently on their children's behavioral problems by completing the Child Behavioral Check List (CBCL/2–3 and CBCL/4–18) at ages 2, 3, 5, 7, and 9 years, (Achenbach, 1991a, 1992). Teachers, who were blind to the children's group affiliations, reported on the Teachers Report Form (TRF) at the ages of 7 and 9 years, (Achenbach, 1991b). Both CBCL and TRF provide standardized measures of child and adolescent emotional/behavioral problems and social competences. Behavioral outcomes are summarized as total problems and broadly scaled on Internalizing and Externalizing dimensions. *Internalizing behavior* consists of the subscales withdrawn, anxious/depressed, and somatic problems, whereas *externalizing behavior* consists of the subscales aggressive and delinquent problems in the CBCL/4–18 and the TRF questionnaires. Behavioral problems were also rated on subscales related to thoughts, social life and attention. Differences between the outcomes of CBCL and TRF take the different environments in which child behavior is observed into account (Achenbach, 1991b, 1992). Parents and teachers also answered questions related to peer problems, prosocial behavior, and the impact supplement extracted from the Strengths and Difficulties Questionnaire (SDQ) at age 9 (Goodman & Scott, 1999). Childbirth and perinatal risk factors were collected before discharge as Clinical Risk Index for Babies (CRIB), existence of bronchopulmonary dysplasia, and existence of several cerebral injuries (Table 1). Parents reported their use of help seeking on behalf of their children at every follow-up. These included type of contact (physiotherapy, child habilitation, child psychiatry services, special educational services, and child welfare authorities), children's age when the contact was initiated, and duration of contact. Social variables (mother's age, years of education, annual income, number of sib-

Table 1
Birth, Medical, and Demographic Characteristics of Infants and Parents

	PI group (n = 72)	PC group (n = 74)	TR group (n = 75)
Infant characteristics			
BW (g), $M \pm SD$	1,396 ± 429	1,381 ± 436	3,619 ± 490
400–1,000 g, n (%)	20 (28)	20 (27)	
1,001–1,500 g, n (%)	15 (21)	20 (27)	
1,501–2,000 g, n (%)	37 (51)	34 (46)	
GA (week), $M \pm SD$	30.2 ± 3.1	29.9 ± 3.5	39.3 ± 1.3
< 28 weeks, n (%)	17 (24)	19 (27)	
28–32 weeks, n (%)	36 (50)	37 (50)	
≥ 33 weeks, n (%)	19 (26)	18 (24)	
Boy, n (%)	38 (53)	39 (53)	40 (54)
Twin, n (%)	16 (22)	14 (19)	0
Parental steroid use, n (%)	53 (74)	57 (77)	
SNAP II, $M \pm SD$	8.3 ± 10.9	10.4 ± 11.3	
CRIB score (N = 85), $M \pm SD$	3.2 ± 2.8	2.7 ± 2.9	
Received ventilation, n (%)	29 (40)	37 (50)	
Duration of ventilation, n (%)	7.0 ± 18.6	7.1 ± 17.3	
Postnatal steroid use, n (%)	9 (13)	10 (14)	
Oxygen therapy at 38 weeks GA, n (%)	11 (15)	14 (19)	
Abnormal cerebral ultrasound, n (%)			
IVH Grade 1 or 2	7 (10)	8 (11)	
IVH Grade 3 or 4	3 (4)	5 (7)	
Periventricular leukomalacia	4 (6)	8 (11)	
Maternal and social characteristics			
Mother's age (years), $M \pm SD$	30.8 ± 6.1	29.1 ± 6.4	29.7 ± 6.1
Firstborn child, n (%)	40 (56)	37 (54)	27 (37)
Mother's education[a] $M \pm SD$	14.6 ± 2.8	13.5 ± 3.2	14.9 ± 2.8
Father's education[a] $M \pm SD$	13.8 ± 3.1	13.5 ± 3.2	14.4 ± 3.2
Mother's monthly income[b]	15.8 ± 7.7	14.6 ± 6.7	15.9 ± 8.0
Father's monthly income[b]	21.1 ± 8.7	19.9 ± 8.1	21.9 ± 9.8

Note. BW = birth weight; CRIB = Clinical Risk Index for Babies; GA = gestational age; IVH = intraventricular hemorrhage; PC = preterm control; PI = preterm intervention; *SD* = standard deviation; SNAP II = Score of Acute Neonatal Physiology II; TR = term reference.
[a]Education in years. [b]In Norwegian 1,000 kroner, calculated for 131 families due to 15 twins.

lings, etc.) were reported by parents on separate questionnaires before discharge from hospital.

Follow-Up Procedures

For the purpose of the study, all participating children received the same medical, developmental, and psychosocial assessments at corrected ages of 6 months and 1, 2, 3, 5, 7, and 9 years. All assessors were blind to the children's group allocation. Questionnaires were collected and families were followed up by a study coordinator at each follow-up session. Parents received written reports about the child's motor and cognitive development after every follow-up. These included recommendations about referrals to other services if relevant, signed by the study director. All results are reported as intention to treat, and the study was approved on three occasions by the regional committee for medical ethics and the Norwegian Data Inspectorate (1999, 2005, and 2010). The study is registered in the ClinicalTrials.gov, NCT00222456.

Analysis

Calculated t scores based on normative samples were used in longitudinal analyses across CBCL/2–3 and CBCL/4–18 as the questionnaires have some different items and different numbers of items. Mean raw scores were used in cross-sectional and descriptive statistics as recommended by Achenbach (1991a, 1991b, 1992). In the longitudinal linear mixed-model analyses, time was treated as a continuous variable. Predicted mean group differences

on each measurement occasion with 95% confidence intervals were also computed. This analysis was still based on a longitudinal model, but in this case time was treated as a categorical variable (Twisk, 2006). By changing the reference time point in the analysis, predicted group differences on each occasion could be estimated. Because of repeated measures and the clustering effects of twin pairs, all the analyses were generated with linear mixed models (LMMs; SPSS statistics, version 20), generating adjusted mean scores for each follow-up age. On Level 1 (within measurements) the covariance structure is the one implied by the standard multilevel model (Singer & Willett, 2003). On Level 2 (within family), a variance components structure was used since the correlation between the slope and the intercept was generally low. Dichotomized variables were analyzed by generalized LMMs, which generated odds ratios adjusted for clustering effects of twin pairs. Effect sizes (ESs) in LMMs were mostly computed as pseudo R^2 and based on the square of the correlation between observed and predicted values of the dependent variable (Singer & Willett, 2003). When assessing the effect of adding variables in a LMM analysis, the change in this pseudo R^2 value is given as an ES measure. For analysis of treatment group comparisons an ES as Hedges' g was computed (Hedges & Olkin, 1985).

Parental agreement was analyzed by intraclass correlations (ICCS) and the difference between the two independent ICC coefficients for the PI and PC groups was tested as described by Alsawalmeh and Feldt (1992).

Results

Participant randomization resulted in well-balanced preterm groups, except for a difference of 1 year in maternal education in favor of the intervention group (Table 1). Thus, maternal education is controlled for in all analyses. In the PI group, fathers with fewer years of education participated in fewer interventional sessions than fathers with more years of education ($F = 4.8$, $p = .03$). Other variables had no significant impact on fathers' participation. Possible implications related to inclusion criteria, outliers, fathers' participation, and parental agreement in the preterm groups will be reported in a final separate paragraph.

The dropout rates were low in all groups throughout the study (Figure 1 and Table 2), and 129 preterms (88%) were still attending the follow-up session at 9 years of age.

Evaluation of Longitudinal Behavioral Development

In responses to the first question, parents did not report different developmental patterns between preterm groups on any of the CBCL problem scales from the age of 2 until 9, nor did teachers report any interactions between groups and age on TRF dimensions from 7 to 9 years. Nonsignificant differences in favor of the PI group were reported (Table 3). The number of problems reported remained at a relatively low level compared to those reported by Achenbach (1991a, 1991b; Figures 2 and 3).

Different patterns were displayed in parents' and teachers' evaluations of internalizing versus externalizing behavior in the preterm groups (Figures 2 and 3). Across groups, all informants reported increased internalizing behavior with age. Teachers also reported increased externalizing behavior, whereas parents reported diminishing levels of problems of externalization. There was considerable variation in parents' and teachers' assessments of the children, as indicated by the high standard deviations in both preterm groups (Table 3).

Table 2
Frequencies of Parents' and Teachers' Reports on Child Behavioral Problems and Competencies

Respondent	PI group (n = 72), n (%)[a]	PC group (n = 74), n (%)[a]	TR group (n = 75), n (%)[a]
2 years			
Mother (CBCL)	62 (86)	59 (80)	63 (84)
Father (CBCL)	57 (79)	49 (66)	59 (79)
3 years			
Mother (CBCL)	69 (96)	66 (89)	67 (89)
Father (CBCL)	60 (83)	53 (72)	58 (77)
5 years			
Mother (CBCL, SDQ[b])	68 (94)	63 (85)	65 (87)
Father (CBCL, SDQ[b])	62 (84)	52 (70)	58 (77)
Preschool teacher (SDQ[b])	60 (83)	54 (73)	59 (79)
7 years			
Mother (CBCL)	68 (94)	64 (87)	63 (84)
Father (CBCL)	56 (78)	53 (72)	55 (73)
Teacher (TRF)	58 (81)	55 (74)	56 (75)
9 years			
Mother (CBCL, SDQ[b])	66 (95)	61 (82)	61 (81)
Father (CBCL, SDQ[b])	55 (76)	50 (68)	53 (71)
Teacher (TRF, SDQ[b])	58 (81)	55 (74)	57 (76)

Note. CBCL = Child Behavior Checklist; PC = preterm control; PI = preterm intervention; SDQ = Strengths and Difficulties Questionnaire; TR = term reference; TRF = Teachers Report Form.
[a]Number of reports and percentage of the original number of participants in the group. [b]Questions extracted from the SDQ.

Table 3
Differences Between the PI and PC Groups as Reported by Mothers and Fathers at 2, 3, 5, 7, and 9 Years of Age on CBCL Main Behavioral Domains

Mean score (*SD*) Adjusted difference, 95% CI[a]	2 years		3 years		5 years		7 years		9 years	
	PC	PI	PC	PI	PC	PI	PC	PI	PC	PI
Total problems										
Mother	31.2 (16.8)	27.4 (15.3)	31.0 (20.1)	28.9 (16.8)	21.6 (16.0)	17.2 (12.6)	22.0 (18.7)	17.6 (16.0)	20.3 (21.2)	17.0 (19.8)
	3.8 [−2.6, 10.2]		1.2 [−4.9, 7.4]		4.2 [−2.0, 10.4]		3.5 [−2.7, 9.7]		2.9 [−3.2, 9.1]	
Father	24.3 (12.2)	22.6 (15.4)	24.8 (13.6)	21.1 (14.6)	17.4 (15.9)	12.5 (8.3)	16.7 (13.4)	10.8 (13.9)	16.8 (25.3)	10.8 (11.2)
	2.9 [−2.9, 8.7]		4.9 [−0.8, 10.6]		5.0 [−0.6, 10.7]		5.4 [−0.3, 11.1]		5.4 [−0.3, 11.2]	
Internalizing										
Mother	7.1 (4.8)	6.2 (4.3)	6.8 (5.1)	5.9 (4.5)	4.5 (4.4)	3.6 (3.5)	5.3 (5.1)	4.6 (4.2)	5.0 (5.4)	4.7 (5.8)
	1.1 [−0.7, 2.8]		0.8 [−0.9, 2.4]		1.0 [−0.6, 2.9]		0.5 [−1.2, 2.1]		0.3 [−1.4, 2.0]	
Father	5.3 (3.4)	4.5 (3.7)	5.4 (4.0)	4.6 (4.1)	2.9 (3.1)	2.9 (2.6)	3.4 (3.7)	2.4 (3.6)	4.1 (7.5)	3.0 (4.0)
	1.0 [−0.6, 2.5]		1.1 [−0.4, 2.7]		0.4 [−1.2, 1.9]		1.0 [−0.5, 2.6]		1.1 [−0.5, 2.7]	
Externalizing										
Mother	10.8 (7.4)	9.3 (5.9)	11.2 (8.2)	10.3 (6.1)	7.9 (6.6)	6.1 (5.4)	7.1 (7.5)	6.0 (6.9)	6.6 (7.8)	5.2 (7.9)
	1.4 [−1.1, 4.0]		0.3 [−2.2, 2.8]		1.3 [−1.2, 3.8]		0.5 [−1.9, 3.0]		0.9 [−1.6, 3.4]	
Father	8.9 (5.5)	7.6 (5.6)	8.9 (5.7)	6.8 (5.1)	6.3 (7.2)	4.4 (3.5)	5.5 (6.2)	3.8 (5.9)	5.2 (9.2)	3.1 (4.0)
	1.6 [−0.7, 3.9]		2.4 [−0.2, 4.6]		1.8 [−0.4, 4.0]		1.3 [−0.9, 3.5]		1.7 [−0.6, 3.9]	

Note. CBCL = Child Behavior Checklist; PC = preterm control; PI = preterm intervention.
[a]Adjusted difference for repeated measures and clustering effects of twin pairs by use of linear mixed models.

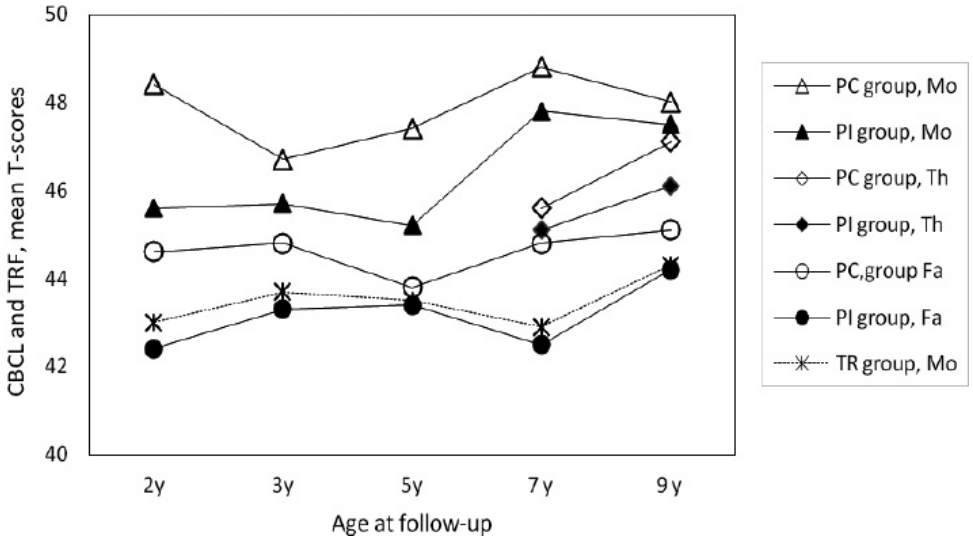

Figure 2. Parents and teachers report on CBCL/TRF, internalizing problems in the PI and PC groups from ages 2 to 9. Because mothers and fathers reported on different questionnaires through childhood (CBCL/2–3 and CBCL/4–18); CBCL = Child *t* scores and not mean scores are presented. A *t* score = 50 represents raw scores at the 50th percentile in a normative American sample (Achenbach, 1991a, 1991b, 1992). PC = preterm control; PI = preterm intervention; TR = term reference; CBCL = Child Behavior Checklist; TRF = Teachers Report Form; Mo = mothers; Fa = Fathers; Th = Teachers.

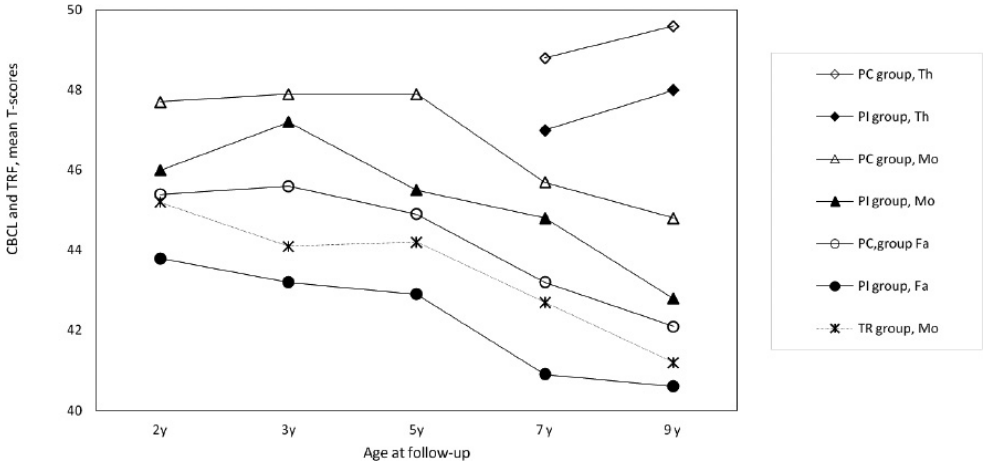

Figure 3. Parents and teachers report on CBCL/TRF, externalizing problems in the PI and PC groups from ages 2 to 9. Because mothers and fathers reported on different questionnaires through childhood (CBCL/2–3 and CBCL/4-18); *t*-scores rather than mean scores are presented. A *t*-score = 50 represents raw scores at the 50th percentile in a normative American sample (Achenbach, 1991a 1991b, 1992). PC = preterm control; PI = preterm intervention; TR = term reference; CBCL = Child Behavior Checklist; TRF = Teachers Report Form; Mo = mothers; Fa = Fathers; Th = Teachers.

Cross-Sectional Differences Between Preterm Groups at Ages 7 and 9

The responses to the second question will be presented in two sections as behavioral problems are reported first and subsequently reports about perceived competencies.

Behavioral Problems

Fewer attention problems in the PI group were reported by mothers, $F(1, 118) = 6.3$, $p = .01$, ES = 0.48; fathers, $F(1, 95) = 4.8$, $p = .03$, ES = 0.48; and teachers, $F(1, 102) = 6.6$, $p = .01$, ES = 0.48, at age 7. Teachers also reported a significant difference

on attention problems at age 9 (see Table 4). At 9 years, both fathers, $F(1, 92) = 8.6$, $p = .004$, ES = 0.52, and teachers (see Table 4) reported lower scores on thought problems in the PI group than in the PC group.

At age 9, teachers perceived fewer difficulties in the PI group than in the PC group, $F(1, 97) = 5.5$, $p = .02$, ES = 0.48. The SDQ question: "Does the child have difficulties?" was subsequently recoded as two categories: (a) absent/minor or (b) definite/severe problems. The PC group had odds that were almost 3 times as high of being perceived as having definite or severe difficulties by teachers at age 9, $t(212) = 2.2$, $p = .03$, OR = 2.9, 95% CI [1.1, 7.6]. Parental evaluations of difficulties corresponded with those of teachers, close to significance for mothers, $F(1, 112) = 3.7$, $p = .058$, ES = 0.39. Fathers in the PI group, who initially had reported their child to have difficulties, perceived more seldom these difficulties as affecting peer relationships compared to fathers in the PC group, $F(1, 31) = 5.5$, $p = .02$, ES = 0.79. A significant correlation between teachers' reports of difficulties and parent's reports regarding referrals to specialized child and adolescent psychiatric services was uncovered, $F(1, 110) = 38.1$, $p < .001$, $r = .57$. Children in the PC group were more than twice as frequently referred as children in the PI group.

Perceived Competences

Generally speaking, in accordance with teachers (Table 4) mothers and fathers perceived a better adaptation to school in the PI group at the ages of 7 and 9 than in the PC group, as reported on the CBCL, total competence scale. At age 9 this generated moderate ESs: mothers, $F(1, 108) = 5.4$, $p = .02$, ES = 0.50, and fathers, $F(1, 84) = 8.6$, $p = .004$, ES = 0.62). A new variable named "school problems" was defined as a score below the 10th percentile for term children of the same sex and age on the variable TRF, academic performance. The PC group experienced significantly more school problems than the PI group at the age of 9, $t(109) = 2.7$, $p = .009$, OR = 3.7, 95% CI [1.4, 9.7].

Outcomes Related to Interactions With Children's BW and Gender

Responses to the third question revealed 1 three-way interaction with group, age, and BW, and 2 three-way interactions with group, age, and gender.

Three-Way Interaction Between Preterm Groups, Age, and BW Groups

Preterm groups were split into subgroups according to whether BW was more or less than

Table 4
Teachers' Reports on Children's Behavioral Problems and Perceived Competences at Age 9 as Reported on TRF

	PI group (n = 55)	PC group (n = 58)	Adjusted mean differences [95% CI]	p	ES[a]	TR group (n = 57)
Reported problems						
Total, *M* (*SD*)	12.4 (13.9)	18.5 (18.5)	−6.1 [−12.2, 0.0]	.06*	0.45	10.4 (16.9)
Internalizing	3.3 (4.5)	3.7 (4.6)	−0.4 [−2.1, 1.2]	.61	0.08	2.1 (3.2)
Externalizing	3.6 (5.5)	5.0 (6.6)	−1.4 [−3.7, 0.8]	.20	0.23	4.3 (7.5)
Social	0.9 (1.7)	1.7 (2.4)	−0.8 [−1.6, −0.1]	.09*	0.38	0.9 (1.9)
Thought	0.2 (0.5)	0.6 (1.3)	−0.4 [−0.7, 0.0]	.05	0.40	0.1 (0.5)
Attention	4.7 (5.4)	7.8 (8.4)	−3.1 [−5.7, −0.5]	.02	0.43	3.1 (5.1)
Aggressiveness	3.1 (4.9)	4.4 (5.9)	−1.3 [−3.3, 0.7]	.20	0.24	3.8 (6.6)
Hyper impulsiveness	2.7 (3.0)	4.3 (4.8)	−1.6 [−3.1, −0.1]	.04	0.39	1.9 (3.4)
Inattention	3.5 (4.2)	5.8 (6.4)	−2.3 [−4.3, −0.3]	.04	0.42	2.2 (3.5)
Perceived competencies						
Academic performance	3.1 (0.5)	2.8 (0.8)	0.3 [0.1, 0.6]	.009	0.45	3.2 (0.5)
Working hard	4.4 (1.3)	3.7 (1.3)	0.7 [0.2, 1.2]	.01	0.54	4.6 (1.2)
Behaving appropriately	4.7 (1.2)	4.1 (1.3)	0.6 [0.2, 1.1]	.01	0.48	4.4 (1.2)
Learning	4.3 (1.3)	3.6 (1.4)	0.7 [0.2, 1.2]	.01	0.52	4.6 (1.1)
Happy	4.4 (1.0)	4.2 (0.8)	0.2 [−0.1, 0.6]	.3	0.22	4.4 (1.0)
Sum of competence items	17.9 (4.0)	15.6 (3.7)	2.3 [0.8, 3.7]	.004	0.60	18.1 (3.6)

Note PC = preterm control; PI = preterm intervention; TR = term reference; TRF = Teachers Report Form.
[a]Effect size (EF) = Hedges' *g*.
*Before adjusting for twin pairs the differences between preterm groups were significant at the .05 level.

1,000 g. A significant three-way interaction between preterm groups, age, and BW on externalizing problems was reported by fathers, $F(1, 417) = 4.5$, $p = .03$. The three-way interaction increased the pseudo R^2 by .029 units. Throughout childhood, fathers reported the heaviest children in the PC group as having more externalizing problems than the other three subgroups (Figure 4).

Three-Way Interaction Between Preterm Groups, Age, and Gender

Mothers reported a three-way interaction between group, age and gender on problems of withdrawal from age 2 until 9, $F(1, 484) = 4.3$, $p = .04$. The three-way interaction increased the pseudo R^2 by .012 units. Mothers of the PC group reported boys as being at a higher level through all preschool ages. However, by age 9 this had fallen to a score close to that of the PI boys. On the other hand, girls were reported to be on the same level in both groups until the age of 7 (within a range of one *t*-score unit). At age 9 the difference between PI and PC girls had increased, as the PI girls were reported to have fewer withdrawn problems than the PC girls.

The second interaction involving group, age, and gender was found in fathers' reports on CBCL anxious/depressed problems from age of 2 until 9, $F(1, 439) = 6.5$, $p = .01$, increased pseudo R^2 by .018. PI and PC fathers reported fewer problems from toddlerhood until age 5 on both boys and girls. From 5 until 9 years of age, the mean score increased for boys in the PC group, while the mean score continued almost unchanged among boys in the PI group. A large part of the increase among the PC boys was due to two persons, which limits the generalizability of this result. Girls in both groups were reported as relatively unchanged from age 5 until 9.

Behavioral Development in Comparison to Term References

No Group × Age interactions between the PI and TR groups were found on any dimensions of behavioral problems. However, one Group × Age interaction was found between the PC and TR group on CBCL, anxious/depressed problems, $F(1, 441) = 9.2$, $p = .003$, increased pseudo R^2 by .013 units, as reported by fathers. The PC and TR groups were scored at similar levels until age 5, but gradually increasing problems were reported at ages 7 and 9 in the PC group.

In cross-sectional analyses significant differences were found on most CBCL behavior dimensions between the PC and TR groups at all follow-ups until age 9. Few significant differences were revealed between the PI and TR groups. At 9 years, no significant differences were reported between the PI and TR groups by mothers, fathers, or teachers.

Results Related to Inclusion Criteria, Outliers, and Respondents

Analyses Related to Inclusion Criteria and Deviant Variable Values

When behavior problems reported on children with disabilities (two in the PI group) were

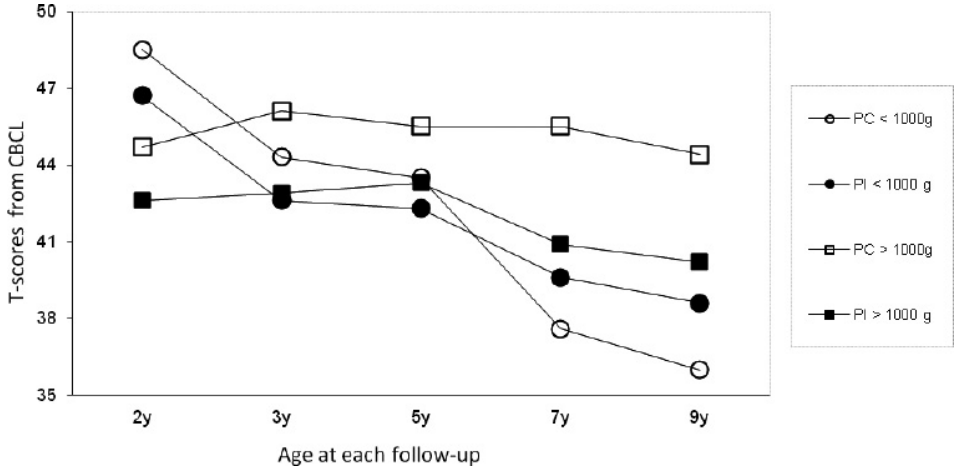

Figure 4. Three-way interaction between group, age, BW, and externalizing problems from age 2 until 9. PC = preterm control; PI = preterm intervention; BW = birth weight; CBCL = Child Behavior Checklist.

included in the analyses, fewer differences between preterm groups remained significant. They were unable to carry out the age-appropriate assessment but the families found it valuable to continue to participate in the study. Teachers' reports of more difficulties in everyday life in the PC group were still evident ($p = .03$).

Data were methodically controlled for potentially disruptive effects of outliers. The consistency was found to be high, with one exception (one child). When these scores were excluded at age 9, several significant differences in favor of the PI group became evident (fewer attention problems and difficulties as reported by mothers, and fewer total and social problems as reported by teachers).

Outcome Variations in the PI Group Related to Fathers' Participation

Fathers participated on average in 6 of 12 sessions (interquartile range = 4–10 sessions). Differences in participation did not influence fathers' reports at age 7. At age 9 fathers who had participated less reported more externalizing problems, $F(1, 98) = 7.5$, $p = .01$, characterized by more aggressive behavior, $F(1, 98) = 6.9$, $p = .01$. These differences were explained by father's length of education, as fathers with relatively less education reported more problems.

Parental Agreement in Ratings of Behavioral Problems

On CBCL main dimensions, the level of agreements between parents in the PC and PI groups changed across childhood. The agreement was higher in the PI group on internalizing scores at age 2 ($ICC_{PC} = 0.26$, $ICC_{PI} = 0.57$, $p = .03$; Alsawalmeh & Feldt, 1992). At age 9 contrary differences were uncovered as agreement between PC parents was higher in reports of externalizing behavior ($ICC_{PC} = 0.76$, $ICC_{PI} = 0.53$, $p = .01$) and total problems ($ICC_{PC} = 0.79$, $ICC_{PI} = 0.63$, $p = .03$).

Discussion

The basic hypothesis in this study was that the intervention could make the immature expressions characterizing preterm infants interpretable for parents and thereby enhance the formation of positive transactions and foster longlasting developmental benefits. Main findings are that the longitudinal trajectory of behavior development, reported on preterm groups until age 9, was not affected by the early intervention even though the intervention may have promoted some protective effects regarding anxious behavior in the PI group. At ages 7 and 9 fewer attentional problems and a better adaptation to school were reported by mothers, fathers, and teachers in the PI group, and similarly, teachers reported a lower incidence of experiencing definite or severe difficulties in everyday life in the PI group. Both preterm groups followed the same trajectory as the TR group, except for a greater likelihood of anxious/depressed behavior in the PC group. Interestingly, at age 9 there were no statistical differences between the PI and the TR group regarding problems or competencies. These findings are consistent with those of the original MITP study (Achenbach et al., 1993). Thus, the basic hypothesis was largely confirmed.

First, similar trajectories of behavior development were reported in the PI and PC groups. From 3 years of age externalizing behavior decreased in both groups while slightly growing internalizing behavior was reported until age 9. Both patterns were consistent with main behavioral dimensions reported in the TR group and normative patterns reported by Bongers et al. (2003). In addition, relatively low levels of problems were reported in TISP, compared to nonintervention studies (Achenbach, 1991a, 1991b; Bongers et al., 2003; Elgen et al., 2002; Hall & Wolke, 2012; Taylor, Klein, Minich, & Hack, 2000). Few children were reported to have more than minor problems in any of the groups. This is consistent with the pattern reported by Hall and Wolke (2012), who found a 75% rate of low emotional problems in a cohort of premature children. As in Elgen et al. (2003), abnormal problems were defined as CBCL, total problems reported above the 90th percentile for the same gender in the term control group. The prevalence of abnormal problems was reported at 15% (mothers and fathers) in the PI group and correspondingly 16% (fathers) and 26% (mothers) in the PC group. Analogous reports from teachers were 11% in the PI group and 20% in the PC group. Elgen et al. (2002) reported abnormal behavioral problems in 40% of their sample and a rate of psychiatric diagnoses at 27% among preterms at 11 years of age (BW < 2,000 g). Comparisons are nevertheless uncertain. We suggest that the low level of problems identified could also be the result of a general strengthening effect caused by the follow-ups, as parents repeatedly received feedback on their child's development.

Internalizing behavior was reported to increase with age across all groups, but parents of preterms generally reported at levels higher than those reported for. This corresponds to reports of more

frequent internalizing and socioemotional problems among preterms than term peers throughout childhood (Aarnoudse-Moens et al., 2009; Loe, Lee, Luna, & Feldman, 2011). On the other hand, gradually diminishing levels of externalizing problems were reported by parents, while teachers reported externalizing problems to rise until age 9. Differences between parents, and teachers, reports may be due to unequal child–adult relationships, environments, and activities where children were observed (Grietens et al., 2004). Bongers et al. (2003) reported that boys displayed more externalizing behavior than girls and a similar tendency was evident across ages in this study, although not statistically significant. In 81% of the studies reviewed by Bhutta et al. (2002), preterm born children were identified as at risk of more externalizing problems than were terms. Other studies have confirmed the falling trend in our data (Aarnoudse-Moens et al., 2009).

Agreement between parents' reports was higher in the PI group compared to the PC group at age 2 (internalizing dimension) but at similar levels across groups from ages 3 to 7. However, at age 9 higher agreements were reported from PC parents (externalizing and total problems) compared to PI parents. Higher agreement is supposed to indicate more shared views (Buehler et al., 1997). In toddlerhood (age 2) the knowledge offered by the intervention may have raised PI parents' sensitivity and endeavoring to interpret children's behavior and feelings, while these challenges may have been more concealed for PC parents. In contrast, at age 9 PC parents reported more challenges among their children and this may have increased their unanimity.

Several cross-sectional differences were identified between the preterm groups at ages 7 and 9. The most pronounced was that the early intervention seems to have strengthened the children's attention. Aarnoudse-Moens et al. (2009) reported increased attentional problems to be a lasting challenge for premature children, ~0.5 *SD* higher than their term peers. Others have confirmed this, showing the connection between prematurity and attentional problems to be mediated by slow responses and deficits in visuospatial working memory (de Kievet, van Elburg, Lafeber, & Oosterlaan, 2012) and impaired cognitive outcomes (Loe et al., 2011). Attentional problems are also thought to be a strong predictor of adaptive problems and academic underachievement (Mulder, Pitchford, Hagger, & Marlow, 2009), and this relates to the next important finding: Both parents and teachers reported the PI group to perform significantly better in school than the PC group. It is promising that the PI group seems to have improved their ability to adapt to school, which is a frequently described major challenge for preterm born children (Aylward, 2005; Bhutta et al., 2002). Significantly more school problems were revealed in the PC group compared to the PI group at age 9. Despite this, at age 9 PC and PI group parents reported similar frequencies of contact with pedagogical psychological services (34% and 28%, respectively). The prevalence is 10 times as high as among the term controls (3%), but moderate compared to about 50% reported by Aylward (2005). The discrepancies between these two findings (similar use of special services but differences in perceived academic performance) may be due to a proactive approach being taken by parents in the PI group, for example, by asking at an early stage for help.

Blind to children's group allocation, teachers of PC children reported more perceived difficulties than the teachers of the PI group at age 9. This result was significantly associated with parents' reports regarding referrals to specialized child and adolescent psychiatric services. Perceived difficulties were reported on the single SDQ question: "Does your child have difficulties?" which has been described as having a strong predictive value for the child's later mental health (Goodman, 1999; Goodman & Goodman, 2011). Children in the PC group were more than twice as often referred as the PI group. This finding may indicate a preventive intervention effect regarding children's mental health problems. The risk of psychiatric disorders among children has been investigated by comparing high scores on parents' ratings of problems. In a study screening for child psychiatric disorders, CBCL total problem scores ≥ 35 were regarded as an appropriate cutoff point and a prevalence of 10.1% among 8- to 9-year-old schoolchildren were reported (Bilenberg, Petersen, Hoerder, & Gillberg, 2005). With a similar cutoff, 18% of the PC group scored in the clinical area while 11.8% in the PI group met this criterion (nonsignificant difference).

Our third question returned to the longitudinal investigation. Analyses revealed two outcomes of behavior where the group allocation may have affected preterm girls and boys differently and one variable where behavior development varied related to the degree of prematurity. First, PC mothers reported their sons closer to PI boys with increasing age with respect to withdrawn behavior, whereas daughters in the PC group were reported

with higher levels of withdrawn behavior at ages 7 and 9 than PI daughters. Second, fathers reported that PC boys displayed more anxious/depressed behavior than PI boys from age 5 to 9, whereas girls were reported at quite similar levels throughout childhood. These two interactions address aspects of internalizing behavior and are of particular interest as higher incidences of anxious and depressed behaviors have repeatedly been reported among premature born children (Johnson & Marlow, 2011; Loe et al., 2011). However, the ESs are small and these interactions may have appeared by chance.

Third, we uncovered different trajectories in fathers' reports of externalizing behavior. Preterms with BW > 1,000 g in the PC group were reported to display more externalizing problems from ages 5 to 9, compared to children with BW > 1,000 g in the PI group and children with BW < 1,000 g in both groups. Previously, less favorable developmental outcomes have been reported among preterms with the lowest weights or GAs (Litt, Taylor, Klein, & Hack, 2005; Saigal et al., 2003). One conceivable reason for the contrasting result may be related to differences in the caretaking offered for fathers in the newborn period. Heavier and medically stable infants stayed in the hospital for shorter periods and their fathers were given limited opportunities to be present and participate in the daily care of their infants (before they were randomized to either the PI or the PC group). Compared to fathers of the most immature children in both groups (with prolonged hospital stay) and PI fathers who participated in the interventions, they received less information and practical guidance on how to understand their baby's cues and expressions. At children's age of 2 the same subgroup of PC fathers reported more child-related stress than the other subgroups and this may have been an early indicator of the same phenomena (Kaaresen et al., 2008). Parenting stress is a strong predictor of both present and later child behavior problems (Gray et al., 2004). At least this result emphasizes the need to ensure that fathers of preterm infants receive adequate amounts of training in seeing, interpreting, and reacting to their infants' cues and expressions.

Finally, we compared behavioral problems reported by PI and PC parents with those reported by parents of term controls. One significant difference in longitudinal development of anxious and depressed behavior was uncovered as fathers reported a more worrisome trajectory of behavior on children in the PC group compared to the TR group. Higher prevalence of internalizing behavior has previously been reported in populations of preterms (Aarnoudse-Moens et al., 2009) and we speculate that the intervention may have offered protective effects as a similar difference not was detected between the PI and TR groups. In addition, significantly more problems were reported in the PC group than in the TR group at all follow-ups. This confirmed the pattern previously described that preterm born children face more problems across childhood compared to term peers (Aarnoudse-Moens et al., 2009; Bhutta et al., 2002). The finding of no cross-sectional differences between the PI and TR groups at age 9 is promising and may indicate that the intervention has contributed to a large degree of normalization related to preterm born children's behavioral development.

Several mechanisms, activated from early childhood, may have contributed to the positive intervention effects found in the PI group at age 9. Olafsen and colleagues (Olafsen et al., 2006; Olafsen et al., 2012) reported that infants in the PI group initiated more social communication at age 1, and in particular enhanced capacity to initiate joint attention was observed among poorly regulated children. Those findings may support the findings of Lawson and Ruff (2004) and Feldman (2009), that a capacity for early focused attention can predict behavioral regulation at age 5. In the first months of life, a period characterized by rapid developmental transitions (Nugent et al., 2007), the infant is completing the change from intrauterine to extrauterine regulation. Schmid et al. (2011) concluded that the amount of maternally initiated contact behavior at a very early developmental stage (assessed at 3 months) may be crucial for children's mental health later in life. An important early regulation of attention and emotion is also taking place (Lavelli & Fogel, 2005). The general focus of the MITP program as designed in the 1980s was to help parents to facilitate and appreciate moments of joint attention with their baby, and thus be able to adjust their own actions so that basic biological rhythms in the child were minimally disturbed (Feldman, 2006). Responsiveness and the quality of face-to-face reciprocity have been described as important contributors to the development of child self-regulation competences, especially in premature children (Evans et al., 2012; Feldman et al., 1999).

Although significant improvements in neonatal care treatments have been introduced since then, similar effects on behavior at age 9 were confirmed in this study. The intervention seems to have had

some unique effects beyond this, due to: (a) the clear structure of parental guidance, (b) introduction of important concepts and understandings about the regulation difficulties of preterms, (c) the enhancement of parental empowerment, and (d) the importance of supporting early social communication between parents and the child. The study by Milgrom et al. (2013) confirms this, as the authors report several positive intervention effects of the MITP as early as at 6 months.

Another important result is the conceit of active involvement and support of fathers. PI fathers, who had participated in relatively few intervention sessions, were less educated and tended to report more externalizing and aggressive behavior at age 9 than those who participated in more interventions. A similar result has been reported by Herbert, Harvey, Lugo-Candelas, and Breaux (2013) as both low paternal socioeconomic status and more paternal depressive symptoms predicted impaired outcomes on a wide range of children's outcomes across early childhood. This highlights the importance of early paternal involvement, especially in less educated families.

Strengths and Limitations

The high participation rate throughout childhood is a major strength of this study. Parents evidently judged the value of the study to be high, and the repeated follow-ups have been a lasting motivational factor. Indeed, the follow-up regime itself could have enhanced parental security. These may have biased our study groups in the direction of lower rates of problems relative to Norwegian general populations of terms and preterms (Jozefiak, Larsson, Wichstrom, & Rimehaug, 2012). Another strength is the study design, whereby preterms were stratified within groups and randomized to intervention or control. The sample size was defined by the aim of finding significant differences between preterm groups at the age of 2 (Kaaresen et al., 2006). The participation rates at 9 years were close to 90%, but the clustering effect of twin pairs in the preterm groups made the sample size somewhat limited. This made analyses of subgroups underpowered and the clinical relevance of results more difficult to interpret.

Other possible limitations of the study are the collection of information about the children's behavior. Parental reports may be highly influenced by their engagement in and knowledge of the study, and these should therefore be treated as subjective information. On the other hand, teachers' reporting on TRF is regarded as being more objective, as teachers were blinded to the children's group affiliation. It should be mentioned that the TR group may not be representative for all terms. However, this group did serve as an important reference.

Further research is needed, as the predischarge part of this early intervention was carried out in the course of each child's last week in the neonatal unit. Similar guidance and parental sensitization begins much earlier in today's clinical NICU practice in many countries. This may affect how the intervention should be designed and implemented, and therefore needs to be investigated. Different versions of the intervention need to be tested following adjustment to the parental presence and family-centered care that is now typical to ensure that the positive intervention effects reported last.

References

Aarnoudse-Moens, C. S. H., Weisglas-Kuperus, N., van Goudoever, J. B., & Oosterlaan, J. (2009). Meta-analysis of neurobehavioral outcomes in very preterm and/or very low-birth-weight children. *Pediatrics*, *124*, 717–728. doi:10.1542/peds.2008-2816

Abidin, R.R. (1995). *Parenting Stress Index (PSI), professional manual (3rd ed.)*. Odessa, FL: Psychological Assessment Resources..

Achenbach, T. (1991a). *Manual for the Child Behavior Checklist/4-18 and 1991 Profile*. Burlington: University of Vermont.

Achenbach, T. M. (1991b). *Manual for the Teacher's Report Form and 1991 Profile*. Burlington: University of Vermont.

Achenbach, T. M. (1992). *Manual for the Child Behavior Checklist/2-3 and 1992 Profile*. Burlington: University of Vermont.

Achenbach, T. M., Howell, C. T., Aoki, M., & Rauh, V. A. (1993). Nine-year outcome of the Vermont Intervention Program for low birth-weight infants. *Pediatrics*, *91*, 45–55.

Alsawalmeh, Y. M., & Feldt, L. S. (1992). Test of the hypothesis that the intraclass reliability coefficient is the same for two measurement procedures. *Applied Psychological Measurement*, *16*, 195–205.

Aylward, G. (2005). Neurodevelopmental outcomes of infants born prematurely. *Journal of Developmental and Behavioral Pediatrics*, *26*, 427–440.

Bayley, M. (1993). *Bayley Scales in Infant Development*. San Antonio, TX: Psychological Cooperation.

Bhutta, A., Cleves, M., Casey, P., Cradock, M., & Anand, K. (2002). Cognitive and behavioral outcomes of school-aged children who were born preterm: A meta-analysis. *Journal of the American Medical Association*, *288*, 728–737.

Bilenberg, N., Petersen, D., Hoerder, K., & Gillberg, C. (2005). The prevalence of child-psychiatric disorders

among 8–9-years-old children in Danish mainstream schools. *Acta Psychiatrica Scandinavica*, *111*, 59–67.

Bongers, I. L., Koot, H. M., van der Ende, J., & Verhulst, F. C. (2003). The normative development of child and adolescent problem behavior. *Journal of Abnormal Psychology*, *112*, 179–192.

Buehler, C., Anthony, C., Krishnakumar, A., Stone, G., Gerard, J., & Pemberton, S. (1997). Interparental conflict and youth problem behaviors: A meta-analysis. *Journal of Child and Family Studies*, *6*, 233–247.

de Kievet, J. F., van Elburg, R. M., Lafeber, H. N., & Oosterlaan, J. (2012). Attention problems of very preterm children compared with age-matched term controls at school-age. *Journal of Pediatrics*, *161*, 824–829. doi:10.1016/j.jpeds.2012.05.010

Elgen, I. B., Holsten, F., & Odberg, M. D. (2012). Psychiatric disorders in low birthweight young adults. Prevalence and association with assessments at 11 years. *European Psychiatry*, *28*, 393–396. doi:10.1016/j.eurpsy.2012.06.002

Elgen, I., & Sommerfelt, K. (2002). Low birthweight children: Coping in school? *Acta Paediatrica Scandinavica*, *91*, 939–945.

Elgen, I., Sommerfelt, K., & Markestad, T. (2002). Population-based, controlled study of behavioural problems and psychiatric disorders in low birthweight children at 11 years of age. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *87*, 128–132.

Evans, T., Whittingham, K., & Boyd, R. (2012). What helps the mother of a preterm infant to become securely attached, responsive and well-adjusted? *Infant Behavior and Development*, *35*, 1–11. doi:10.1016/j.infbeh.2011.10.002

Feldman, R. (2006). From biological rhythms to social rhythms: Physiological precursors of mother-infant synchrony. *Developmental Psychology*, *42*, 175–188.

Feldman, R. (2009). The development of regulatory functions from birth to 5 years: Insights from premature infants. *Child Development*, *80*, 544–561.

Feldman, R., Greenbaum, C. W., & Yirmiya, N. (1999). Mother–infant affect synchrony as an antecedent of the emergence of self-control. *Developmental Psychology*, *35*, 223–231.

Glazebrook, C., Marlow, N., Israel, C., Croudace, T., Johnson, S., White, I. R., & Whitelaw, A. (2007). Randomised trial of a parenting intervention during neonatal intensive care. *Archives of Disease in Childhood: Fetal and Neonatal Edition*, *92*, F438–F443.

Goodman, A., & Goodman, R. (2011). Population mean scores predict child mental disorder rates: Validating SDQ prevalence estimators in Britain. *Journal of Child Psychology and Psychiatry*, *52*, 100–108. doi:10.1111/j.1469-7610.2010.02278.x

Goodman, R. (1999). The extended version of the Strengths and Difficulties Questionnaire as a guide to child psychiatric caseness and consequent burden. *Journal of Child Psychology and Psychiatry*, *40*, 791–799.

Goodman, R., & Scott, S. (1999). Comparing the Strengths and Difficulties Questionnaire and the Child Behavior Checklist: Is small beautiful? *Journal of Abnormal Psychology*, *27*, 17–24.

Gray, R. F., Indurkhya, A., & McCormick, M. C. (2004). Prevalence, stability, and predictors of clinically significant behavior problems in low birth weight children at 3, 5, and 8 years of age. *Pediatrics*, *114*, 736–743.

Grietens, H., Onghena, P., Prinzie, P., Gadeyne, E., van Assche, V., Ghesquière, P., & Hellinckx, W. (2004). Comparison of mothers', fathers', and teachers' reports on problems behavior in 5- to 6-years-old children. *Journal of Psychopathology and Behavioral Assessment*, *26*, 137–145.

Guralnick, M. J. P. (2012). Preventive interventions for preterm children: Effectiveness and developmental mechanisms. *Journal of Developmental and Behavioral Pediatrics*, *33*, 353–364. doi:10.1097/DBP.0b013e31824eaa3c

Hall, J., & Wolke, D. (2012). A comparison of prematurity and small for gestational age as risk factors for age 6–13 years emotional problems. *Early Human Development*, *88*, 797–804. doi:10.1016/j.earlhumdev.2012.05.005

Hauser-Cram, P., Warfield, M. E., Shonkoff, J. P., Krauss, M. W., Sayer, A., Upshur, C. C., & Hodapp, R. M. (2001). Children with disabilities: A longitudinal study of child development and parent well-being. *Monographs of the Society for Research in Child Development*, *66*(3, Serial No. 266).

Hedges, L. V., & Olkin, L. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Herbert, S. D., Harvey, E. A., Lugo-Candelas, C. I., & Breaux, R. P. (2013). Early fathering as a predictor of later psychosocial functioning among preschool children with behavior problems. *Journal of Abnormal Psychology*, *41*, 691–703. doi:10.1007/s10802-012-9706-8

Johnson, S. (2007). Cognitive and behavioural outcomes following very preterm birth. *Seminars in Fetal & Neonatal Medicine*, *12*, 363–373.

Johnson, S., & Marlow, N. (2011). Preterm birth and childhood psychiatric disorders. *Pediatric Research*, *69*, 11R–18R. doi:10.1203/PDR.0b013e318212faa0

Jozefiak, T., Larsson, B., Wichstrom, L., & Rimehaug, T. (2012). Competence and emotional/behavioural problems in 7–16-years-old Norwegian school children as reported by parents. *Nordic Journal of Psychiatry*, *66*, 311–319.

Kaaresen, P. I., Rønning, J. A., Tunby, J., Nordhov, M. S., Ulvund, S. E., & Dahl, D. B. (2008). A randomized controlled trial of an early intervention program in low birth weight children: Outcomes at 2 years. *Early Human Development*, *84*, 201–209. doi:10.1016/j.earlhumdev.2007.07.003

Kaaresen, P. I., Rønning, J. A., Ulvund, S. E., & Dahl, L. B. (2006). A randomized, controlled trial of the effectiveness of an early-intervention program in reducing parenting stress after preterm birth. *Pediatrics*, *118*, e9–e19.

Lavelli, M., & Fogel, A. (2005). Developmental changes in the relationship between the infant's attention and emotion during early face-to-face communication: The 2-month transition. *Developmental Psychology*, *41*, 265–280. doi:10.1037/0012-1649.41.1.265

Lawson, K., & Ruff, H. (2004). Early focused attention predicts outcome for children born prematurely. *Journal of Developmental and Behavioral Pediatrics*, *25*, 399–406.

Litt, J., Taylor, G., Klein, N., & Hack, M. (2005). Learning disabilities in children with very low birthweight: Prevalence, neuropsychological correlations and educational interventions. *Journal of Learning Disabilities*, *38*, 130–141.

Loe, I. M., Lee, E. S., Luna, B., & Feldman, H. M. (2011). Behavior problems of 9–16 years old preterm children: Biological, sociodemographic, and intellectual contributions. *Early Human Development*, *87*, 247–252. doi:10.1016/j.earlhumdev.2011.01.023

McAnulty, G. B., Duffy, F. H., Butler, S. C., Bernstein, J. H., Zurakowski, D., & Als, H. (2010). Effects of the Newborn Individualized Developmental Care and Assessment Program (NIDCAP) at age 8 years: Preliminary data. *Clinical Pediatrics*, *49*, 258–270. doi:10.1177/0009922809335668

McCarton, C., Brooks-Gunn, J., Wallace, I., Bauer, C., Bennett, F., Bernbaum, J., & Meinert, C. (1997). Results at age 8 years of early intervention for low-birth-weight premature infants. The Infant Health and Development Program. *Journal of the American Medical Association*, *277*, 126–132.

Meijssen, D., Wolf, M. J., Koldewijn, K., Houtzager, B. A., van Wassenaer, A., Tronick, E., & van Baar, A. (2010). The effect of the infant behavioral assessment and intervention program on mother-infant interaction after very preterm birth. *Journal of Child Psychology and Psychiatry*, *51*, 1287–1295. doi:10.1111/j.1469-7610.2010.02237.x

Milgrom, J., Newnham, C., Martin, P. R., Anderson, P. J., Doyle, L. W., Hunt, R. W., & Gemmil, A. W. (2013). Early communication in preterm infants following intervention in the NICU. *Early Human Development*, *89*, 755–762.

Mulder, H., Pitchford, N. J., Hagger, M. S., & Marlow, N. (2009). Development of executive function and attention in preterm children: A systematic review. *Developmental Neuropsychology*, *34*, 393–421. doi:10.1080/87565640902964524

Newnham, C. A., Milgrom, J., & Skouteris, H. (2009). Effectiveness of a modified mother-infant transaction program on outcomes for preterm infants from 3 to 24 months of age. *Infant Behavior and Development*, *32*, 17–26. doi:10.1016/j.infbeh.2008.09.004

Nordhov, S. M., Rønning, J. A., Dahl, L. B., Ulvund, S. E., Tunby, J., & Kaaresen, P. I. (2010). Early intervention improves cognitive outcomes for preterm infants: A randomized controlled trial. *Pediatrics*, *126*, e1088–e1094.

Nordhov, S. M., Rønning, J. A., Ulvund, S. E., Dahl, L. B., & Kaaresen, P. I. (2012). Early intervention improves behavioral outcomes for preterm infants: Randomized controlled trial. *Pediatrics*, *129*, e9–e16. doi:10.1542/peds.2011-0248

Nugent, K. J., Keefer, C. H., Minear, S., Johnson, L. C., & Blanchard, Y. (2007). *Understanding newborn behavior relationships*. Baltimore, MD: Paul H. Brookes Publishing Co.

Nurcombe, B., Howell, D. C., Rauh, V. A., Teti, D. M., Ruoff, P., & Brennan, J. (1984). An intervention for mothers of low-birthweight infants: Preliminary results. *Journal of the American Academy of Child Psychiatry*, *23*, 319–325.

Olafsen, K. S., Kaaresen, P. I., Handegard, B. H., Ulvund, S. E., Dahl, L. B., & Rønning, J. A. (2008). Maternal ratings of infant regulatory competence from 6 to 12 months: Influence of perceived stress, birth-weight, and intervention: A randomized controlled trial. *Infant Behavior and Development*, *31*, 408–421. doi:10.1016/j.infbeh.2007.12.005

Olafsen, K. S., Rønning, J. A., Handegard, B. H., Ulvund, S. E., Dahl, L. B., & Kaaresen, P. I. (2012). Regulatory competence and social communication in term and preterm infants at 12 months corrected age. Results from a randomized controlled trial. *Infant Behavior and Development*, *35*, 140–149. doi:10.1016/j.infbeh.2011.08.001

Olafsen, K. S., Rønning, J. A., Kaaresen, P. I., Ulvund, S. E., Handegard, B. H., & Dahl, L. B. (2006). Joint attention in term and preterm infants at 12 months corrected age: The significance of gender and intervention based on a randomized controlled trial. *Infant Behavior and Development*, *29*, 554–563. doi:10.1016/j.infbeh.2006.07.004

Rauh, V. A., Achenbach, T. M., Nurcombe, B., Howell, C. T., & Teti, D. M. (1988). Minimizing adverse effects of low birthweight: Four-years results of an early intervention. *Child Development*, *59*, 544–553.

Rauh, V. A., Nurcombe, B., Achenbach, T., & Howell, C. (1990). The Mother–Infant Transaction Program. The content and implications of an intervention for the mothers of low-birthweight infants. *Clinics in Perinatology*, *17*, 31–45.

Rønning, J. A., Dahl, L. B., Ulvund, S. E., & Kaaresen, P. I. (1999). *Can early intervention prevent developmental disturbances/delays and improve the health of children born prematurely? (Study protocol)*. Department of Child And Adolescent Psychiatry: University of Tromsø, Tromsø, Norway.

Saigal, S., Ouden, L. D., Wolke, D., Hoult, L., Streiner, D. L., Withaker, A., & Pinto-Martin, J. (2003). School-age outcomes in children who were extremely low birth weight from four international population-based cohorts. *Pediatrics*, *112*, 943–950. doi:10.1542/peds.112.4.943

Sameroff, A. J., & Fiese, B. (2000). Transactional regulation: The developmental ecology of early intervention. In J. Shonkoff & S. Meisel (Eds.), *Handbook of early childhood intervention* (2nd ed., pp. 135–159). Cambridge, UK: Cambridge University Press. Available from http://dx.doi.org/10.1017/CBO9780511529320.009

Schmid, B., Blomeyer, D., Buchmann, A., Trautmann-Villalba, P., Zimmermann, U., Schmidt, M., . . . Laucht M. (2011). Quality of early mother–child interaction associated with depressive psychopathology in the offspring: A prospective study from infancy to adulthood. *Journal of Psychiatric Research*, *45*, 1387–1394.

Shonkoff, J., & Phillips, D. (Eds.). (2000). *From neurons to neighborhoods: The science of early childhood development*. Washington, DC: National Academy Press.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis, modeling change and event occurrence*. New York, NY: Oxford University Press.

Taylor, G. H., Klein, N., Minich, N. M., & Hack, M. (2000). Middle-school-age outcomes in children with very low birthweight. *Child Development*, *71*, 1495–1511.

Twisk, J. W. R. (2006). *Applied multilevel analysis*. Cambridge, UK: Cambridge University Press.