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, Craddock, & Anand, 2002). A higher incidence of attentional deficits is found to be the most obvious problem, but internalizing problems 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, Indurkhy, & 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., 2008; Song, 2008). 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 Haier, 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.
preterms is also known to in
2007). (Nugent, Keefer, Minear, Johnson, & Blanchard, 2006). This limits the possibilities of caregivers
and emit recognizable alertness responses (Feld-
mance, 2006). 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; Nord-
hov, Rønning, Ulvund, Dahl, & Kaaresen, 2012;
Olafsen et al., 2008; Olafsen et al., 2012). In the
Vermont study, Rauh, Achenbach, Nurcombe, Ho-
well, and Teti (1988) suggested that the interven-
tion 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 interven-
tion effects on behavioral development of preterm
children until age 9. Both longitudinal and cross-
sectional investigations were incorporated in the

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 reinforce-
ment 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 (physio-
logical, 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

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 inter-
human environments. This is formed by the context-
tual interplay between individual biological
conditions (child) and the child’s interactions with
the environment (family, school). As stated by Ha-
user-Cram et al. (2001), “Development occurs
within multiple hierarchical contexts that are recip-
rocally related” (p. 21). When children are born pre-
maturely, the natural transactions are easily
interrupted because of these infants’ immature
expressions and limited capacity to self-regulate
and emit recognizable alertness responses (Feld-
man, 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
preterm children is also known to influence development
(Barlow, 1987; 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.

This study investigated the long-term interven-
tion 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
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-

---

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

### Table 1

**Birth, Medical, and Demographic Characteristics of Infants and Parents**

<table>
<thead>
<tr>
<th></th>
<th>PI group (n = 72)</th>
<th>PC group (n = 74)</th>
<th>TR group (n = 75)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Infant characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BW (g), M ± SD</td>
<td>1,396 ± 429</td>
<td>1,381 ± 436</td>
<td>3,619 ± 490</td>
</tr>
<tr>
<td>400–1,000 g, n (%)</td>
<td>20 (28)</td>
<td>20 (27)</td>
<td></td>
</tr>
<tr>
<td>1,001–1,500 g, n (%)</td>
<td>15 (21)</td>
<td>20 (27)</td>
<td></td>
</tr>
<tr>
<td>1,501–2,000 g, n (%)</td>
<td>37 (51)</td>
<td>34 (46)</td>
<td></td>
</tr>
<tr>
<td>GA (week), M ± SD</td>
<td>30.2 ± 3.1</td>
<td>29.9 ± 3.5</td>
<td>39.3 ± 1.3</td>
</tr>
<tr>
<td>&lt; 28 weeks, n (%)</td>
<td>17 (24)</td>
<td>19 (27)</td>
<td></td>
</tr>
<tr>
<td>28–32 weeks, n (%)</td>
<td>36 (50)</td>
<td>37 (50)</td>
<td></td>
</tr>
<tr>
<td>≥ 33 weeks, n (%)</td>
<td>19 (26)</td>
<td>18 (24)</td>
<td></td>
</tr>
<tr>
<td>Boy, n (%)</td>
<td>38 (53)</td>
<td>39 (53)</td>
<td>40 (54)</td>
</tr>
<tr>
<td>Twin, n (%)</td>
<td>16 (22)</td>
<td>14 (19)</td>
<td>0</td>
</tr>
<tr>
<td>Parental steroid use, n (%)</td>
<td>53 (74)</td>
<td>57 (77)</td>
<td></td>
</tr>
<tr>
<td>SNAP II, M ± SD</td>
<td>8.3 ± 10.9</td>
<td>10.4 ± 11.3</td>
<td></td>
</tr>
<tr>
<td>CRIB score (N = 85), M ± SD</td>
<td>3.2 ± 2.8</td>
<td>2.7 ± 2.9</td>
<td></td>
</tr>
<tr>
<td>Received ventilation, n (%)</td>
<td>29 (40)</td>
<td>37 (50)</td>
<td></td>
</tr>
<tr>
<td>Duration of ventilation, n (%)</td>
<td>7.0 ± 18.6</td>
<td>7.1 ± 17.3</td>
<td></td>
</tr>
<tr>
<td>Postnatal steroid use, n (%)</td>
<td>9 (13)</td>
<td>10 (14)</td>
<td></td>
</tr>
<tr>
<td>Oxygen therapy at 38 weeks GA, n (%)</td>
<td>11 (15)</td>
<td>14 (19)</td>
<td></td>
</tr>
<tr>
<td>Abnormal cerebral ultrasound, n (%)</td>
<td>7 (10)</td>
<td>8 (11)</td>
<td></td>
</tr>
<tr>
<td>IVH Grade 1 or 2</td>
<td>3 (4)</td>
<td>5 (7)</td>
<td></td>
</tr>
<tr>
<td>Periventricular leukomalacia</td>
<td>4 (6)</td>
<td>8 (11)</td>
<td></td>
</tr>
<tr>
<td><strong>Maternal and social characteristics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother’s age (years), M ± SD</td>
<td>30.8 ± 6.1</td>
<td>29.1 ± 6.4</td>
<td>29.7 ± 6.1</td>
</tr>
<tr>
<td>Firstborn child, n (%)</td>
<td>40 (56)</td>
<td>37 (54)</td>
<td>27 (37)</td>
</tr>
<tr>
<td>Father’s education&lt;sup&gt;a&lt;/sup&gt; M ± SD</td>
<td>14.6 ± 2.8</td>
<td>13.5 ± 3.2</td>
<td>14.9 ± 2.8</td>
</tr>
<tr>
<td>Mother’s monthly income&lt;sup&gt;b&lt;/sup&gt;</td>
<td>13.8 ± 3.1</td>
<td>13.5 ± 3.2</td>
<td>14.4 ± 3.2</td>
</tr>
<tr>
<td>Father’s monthly income&lt;sup&gt;b&lt;/sup&gt;</td>
<td>15.8 ± 7.7</td>
<td>14.6 ± 6.7</td>
<td>15.9 ± 8.0</td>
</tr>
</tbody>
</table>
| **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.  
<sup>a</sup>Education in years.  
<sup>b</sup>In Norwegian 1,000 kroner, calculated for 131 families due to 15 twins.
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>
<thead>
<tr>
<th></th>
<th>PI group ($n = 72$)</th>
<th>PC group ($n = 74$)</th>
<th>TR group ($n = 75$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mother (CBCL)</td>
<td>62 (86)</td>
<td>59 (81)</td>
<td>63 (84)</td>
</tr>
<tr>
<td>Father (CBCL)</td>
<td>57 (79)</td>
<td>49 (66)</td>
<td>59 (79)</td>
</tr>
<tr>
<td>3 years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother (CBCL)</td>
<td>69 (96)</td>
<td>66 (89)</td>
<td>67 (89)</td>
</tr>
<tr>
<td>Father (CBCL)</td>
<td>60 (83)</td>
<td>53 (72)</td>
<td>58 (77)</td>
</tr>
<tr>
<td>5 years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother (CBCL, SDQ)</td>
<td>68 (94)</td>
<td>63 (85)</td>
<td>65 (87)</td>
</tr>
<tr>
<td>Father (CBCL, SDQ)</td>
<td>62 (84)</td>
<td>52 (70)</td>
<td>58 (77)</td>
</tr>
<tr>
<td>Preschool teacher (SDQ)</td>
<td>60 (83)</td>
<td>54 (73)</td>
<td>59 (79)</td>
</tr>
<tr>
<td>7 years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother (CBCL)</td>
<td>68 (94)</td>
<td>64 (87)</td>
<td>63 (84)</td>
</tr>
<tr>
<td>Father (CBCL)</td>
<td>56 (78)</td>
<td>53 (72)</td>
<td>55 (73)</td>
</tr>
<tr>
<td>Teacher (TRF)</td>
<td>58 (81)</td>
<td>55 (74)</td>
<td>56 (75)</td>
</tr>
<tr>
<td>9 years</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother (CBCL, SDQ)</td>
<td>66 (95)</td>
<td>61 (82)</td>
<td>61 (81)</td>
</tr>
<tr>
<td>Father (CBCL, SDQ)</td>
<td>55 (76)</td>
<td>50 (68)</td>
<td>53 (71)</td>
</tr>
<tr>
<td>Teacher (TRF, SDQ)</td>
<td>58 (81)</td>
<td>55 (74)</td>
<td>57 (76)</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Mean score (SD)</th>
<th>Adjusted difference, 95% CI$^a$</th>
<th>2 years</th>
<th>3 years</th>
<th>5 years</th>
<th>7 years</th>
<th>9 years</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PC</td>
<td>PI</td>
<td>PC</td>
<td>PI</td>
<td>PC</td>
<td>PI</td>
</tr>
<tr>
<td><strong>Total problems</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>31.2 (16.8)</td>
<td>27.4 (15.3)</td>
<td>31.0 (20.1)</td>
<td>28.9 (16.8)</td>
<td>21.6 (16.0)</td>
<td>17.2 (12.6)</td>
</tr>
<tr>
<td>Father</td>
<td>24.3 (12.2)</td>
<td>22.6 (15.4)</td>
<td>24.8 (13.6)</td>
<td>21.1 (14.6)</td>
<td>17.4 (15.9)</td>
<td>12.5 (8.3)</td>
</tr>
<tr>
<td><strong>Internalizing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>7.1 (4.8)</td>
<td>6.2 (4.3)</td>
<td>6.8 (5.1)</td>
<td>5.9 (4.5)</td>
<td>4.5 (4.4)</td>
<td>3.6 (3.5)</td>
</tr>
<tr>
<td>Father</td>
<td>5.3 (3.4)</td>
<td>4.5 (3.7)</td>
<td>5.4 (4.0)</td>
<td>4.6 (4.1)</td>
<td>2.9 (3.1)</td>
<td>2.9 (2.6)</td>
</tr>
<tr>
<td><strong>Externalizing</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mother</td>
<td>10.8 (7.4)</td>
<td>9.3 (5.9)</td>
<td>11.2 (8.2)</td>
<td>10.3 (6.1)</td>
<td>7.9 (6.6)</td>
<td>6.1 (5.4)</td>
</tr>
<tr>
<td>Father</td>
<td>8.9 (5.5)</td>
<td>7.6 (5.6)</td>
<td>8.9 (5.7)</td>
<td>6.8 (5.1)</td>
<td>6.3 (7.2)</td>
<td>4.4 (3.5)</td>
</tr>
</tbody>
</table>

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.
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.

Table 4

<table>
<thead>
<tr>
<th>Reported problems</th>
<th>PI group ($n = 55$)</th>
<th>PC group ($n = 58$)</th>
<th>Adjusted mean differences [95% CI]</th>
<th>$p$</th>
<th>ES*</th>
<th>TR group ($n = 57$)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Perceived competencies</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Academic performance</td>
<td>3.1 (0.5)</td>
<td>2.8 (0.8)</td>
<td>0.3 [0.1, 0.6]</td>
<td>.009</td>
<td>0.45</td>
<td>3.2 (0.5)</td>
</tr>
<tr>
<td>Working hard</td>
<td>4.4 (1.3)</td>
<td>3.7 (1.3)</td>
<td>0.7 [0.2, 1.2]</td>
<td>.01</td>
<td>0.54</td>
<td>4.6 (1.2)</td>
</tr>
<tr>
<td>Behaving appropriately</td>
<td>4.7 (1.2)</td>
<td>4.1 (1.3)</td>
<td>0.6 [0.2, 1.1]</td>
<td>.01</td>
<td>0.48</td>
<td>4.4 (1.2)</td>
</tr>
<tr>
<td>Learning</td>
<td>4.3 (1.3)</td>
<td>3.6 (1.4)</td>
<td>0.7 [0.2, 1.2]</td>
<td>.01</td>
<td>0.52</td>
<td>4.6 (1.1)</td>
</tr>
<tr>
<td>Happy</td>
<td>4.4 (1.0)</td>
<td>4.2 (0.8)</td>
<td>0.2 [−0.1, 0.6]</td>
<td>.3</td>
<td>0.22</td>
<td>4.4 (1.0)</td>
</tr>
<tr>
<td>Sum of competence items</td>
<td>17.9 (4.0)</td>
<td>15.6 (3.7)</td>
<td>2.3 [0.8, 3.7]</td>
<td>.004</td>
<td>0.60</td>
<td>18.1 (3.6)</td>
</tr>
</tbody>
</table>

Note: PC = preterm control; PI = preterm intervention; TR = term reference; TRF = Teachers Report Form.

*Effect size (ES) = Hedges’ $g$.

*Before adjusting for twin pairs the differences between preterm groups were significant at the .05 level.

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
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 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 $\times$ Age interactions between the PI and TR groups were found on any dimensions of behavioral problems. However, one Group $\times$ 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...
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 ($\text{ICC}_{\text{PC}} = 0.26, \text{ICC}_{\text{PI}} = 0.57, p = .03$; Alsawalmeh & Feldt, 1992). At age 9 contrary differences were uncovered as agreement between PC parents was higher in reports of externalizing behavior ($\text{ICC}_{\text{PC}} = 0.76, \text{ICC}_{\text{PI}} = 0.53, p = .01$) and total problems ($\text{ICC}_{\text{PC}} = 0.79, \text{ICC}_{\text{PI}} = 0.63, p = .03$).

*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 PI 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, Lefeber, & 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 extraterine 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 MITTP as early as at 6 months.

Another important result is the concept 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 more participation 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.


diagnosis between parents and the child. The study by Achenbach, T. M. (1992).

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


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


