

Long-Term Clinical Significance of the Prevention Programme for Externalizing Problem Behaviour (PEP)

Charlotte Hanisch^a Christopher Hautmann^b Ilka Eichelberger^b Julia Plück^b Manfred Döpfner^b

^a Department of Social and Cultural Sciences, Dusseldorf University of Applied Sciences,

^b Hospital of Psychiatry and Psychotherapy for Children and Adolescents, University of Cologne, Germany

Keywords

Prevention · Externalizing problem behaviour ·
Preschool · Clinical significance

Summary

Background: Behavioural parent training effectively improves child disruptive behavioural problems in preschoolers by increasing parenting competence. Intervention effects are generally assessed by means of statistical group comparisons. The indicated Prevention Programme for Externalizing Problem behaviour (PEP) was shown to effectively reduce child problem behaviour in a control group as well as in an application study. The present analysis investigates the long-term clinical significance of these effects. **Materials and Methods:** A screening instrument was used to generate an indicated sample. The control group (CG) comprised 34 families. In 21 cases, only teachers participated in the PEP training (T) and in 38 families, teachers and parents (P+T) were trained. Child problem behaviour was assessed before and up to 30 months after the training. Clinical relevance of the training effects was evaluated using a measure of clinical significance recommended by Jacobson and Truax. **Results:** Repeated-measures analysis of variance did not yield significant group-by-time interaction effects. Clinically significant improvements, however, were present in 34.2% of the children of the P+K group. 28.6% children of the T group improved significantly, whereas in the CG 17.6% or 32.4% (subject to the assessment method) showed significant clinical improvement. In the P+T group, improvements were already found directly post intervention (36.8% and 39.5%, depending on the assessment instrument), whereas in the CG only 20.6% improved significantly from pre to post. The P+T group showed 42.9% improvement from pre to post. **Conclusions:** In the combined intervention group relevant improvements of problem behaviours were found earlier, suggesting that PEP helps children reduce problem behaviour earlier than in the control group.

Schlüsselwörter

Prävention · Expansives Problemverhalten ·
Kindergartenalter · Klinische Signifikanz

Zusammenfassung

Hintergrund: Lerntheoretisch orientierte Elterntrainings gelten als effektive Präventions- und Interventionsmethode zur Reduzierung expansiver Verhaltensauffälligkeiten im Kindesalter. Interventionseffekte werden hierbei meist im statistischen Gruppenvergleich überprüft. Unser Präventionsprogramm für Expansives Problemverhalten (PEP) war sowohl in einer Kontrollgruppenstudie als auch in einer Anwendungsstudie in der Lage, kindliches Problemverhalten zu reduzieren. Die vorliegende Analyse überprüft die langfristige klinische Signifikanz dieser Veränderungen. **Material und Methoden:** Mithilfe eines Screeninginstruments wurde eine Stichprobe expansiv auffälliger Kindergartenkinder identifiziert. In einer Kontrollgruppe (KG: n = 34), einer Erziehertrainingsgruppe (ER: n = 21) und einer Eltern- und Erzieher-Trainingsgruppe (EL+ER: n = 38) wurde das kindliche Problemverhalten vor und bis 30 Monate nach dem PEP-Training erhoben. Die klinische Relevanz der Trainingseffekte wurde anhand des von Jacobson und Truax vorgeschlagenen Maßes für klinische Signifikanz überprüft. **Ergebnisse:** In einer Messwiederholungsvarianzanalyse zeigten sich keine gruppenspezifischen Verläufe im Untersuchungszeitraum. Klinisch signifikante Verbesserungen erzielten in der kombinierten EL+ER-Gruppe 34,2% der Kinder, in der ER-Gruppe 28,6% und in der KG je nach Instrument 17,6% bzw. 32,4%. In der EL+ER-Gruppe waren diese Verbesserungen bereits unmittelbar nach Interventionsende vorhanden (36,8% bzw. 39,5%), während sich in der KG von Prä nach Post lediglich 20,6% klinisch signifikant verbesserten. Die ER-Gruppe weist zum Postzeitpunkt 42,9% klinisch relevant gebesserte Kinder auf. **Schlussfolgerungen:** Das Problemverhalten wird in der kombinierten Interventionsgruppe früher relevant reduziert als in der Kontrollgruppe. PEP scheint somit in der Lage, eine Reduktion von Problemverhalten, die auch in der Kontrollgruppe zu beobachten ist, früher anzustoßen.

Introduction

Externalizing behaviour disorders – i.e., hyperkinetic, oppositional, and aggressive conduct – are widespread in children of preschool and primary school age [Döpfner et al., 2008a; Kuschel et al., 2004], and pose a high risk of poor social, emotional, educational and professional development [Biederman et al., 2006; Elkins et al., 2007; Fontaine et al., 2008]. The persistence of these behavioural disorders into adulthood poses the need for effective methods of prevention and intervention [Bongers et al., 2004; Tremblay, 2006].

Parenting plays an important role in the emergence and perpetuation of externalizing behaviour disorders [Patterson et al., 2004]. Training sessions based on social learning theories are therefore widely used to promote parenting skills [Lundahl et al., 2006]. These are effective methods of intervention both for clinically significant externalizing disorders and for prevention of externalizing behaviour disorders, with small to moderate effect sizes [Dretzke et al., 2005; Eyberg et al., 2008; Kaminski et al., 2008].

The available prevention programmes are directed either to all interested parents (universal prevention), to families with special risk factors (selective prevention) or to parents whose children have shown the first signs of disorder (indicated prevention).

The indicated Prevention Programme for Externalizing Problem Behaviour (PEP) was developed on the basis of the Therapieprogramm für Kinder mit hyperkinetischem und oppositionellem Problemverhalten (THOP, Therapeutic Programme for Children with Hyperkinetic and Oppositional Problem Behaviour) [Döpfner et al., 2007], which has proven to be effective in long-term symptom reduction [Döpfner et al., 2004]. It provides parents and teachers of preschool children who have externalizing problem behaviour, with behaviour modification techniques and strategies to improve their relationship with the child [Plück et al., 2006].

PEP has been evaluated in a randomised controlled trial of an indicated sample. After the training programme, there were improvements of the child's problem behaviour, of parenting skills and of parent-child interaction [Hanisch et al., 2006], and these effects were stronger if the parents had attended more than half of the training sessions [Hanisch et al., 2010]. In a subsequent application study, staff members of counselling centres and early learning centres were trained to lead PEP parent groups during routine care. The participants in the PEP parent groups were parents of elementary and preschool children with externalizing problem behaviour, who had been involved with these facilities on an ongoing basis. After the programme, they reported an increase in their parenting skills and a decrease in the child's problem behaviour [Hautmann et al., 2008, 2009b]; the most affected children benefited most from the preventive programme [Hautmann et al., 2010]. These effects proved stable after 1 year [Hautmann et al., 2009a].

So far, only a few studies have also considered the clinical significance of the changes achieved by parent training [Nixon et al., 2003]. In this context, a measure is understood as clinically significant if it considers the effects of training on the level of the individual and of clinical practice [Kirk, 2001]. Ogles and colleagues [2001] distinguish three quantitative operationalisations of clinical significance: (1) The change from a clinically relevant disorder to a normal level of functioning, so that, for example, the person no longer differs from healthy subjects after an intervention. (2) A statistically reliable individual change (reliable change index [RCI]), i.e., the change measured on an individual basis goes beyond what could be explained by a measurement error. (3) A combination of the two criteria, i.e., a return to normal function, provided that a change has been demonstrated that is clearly greater than could be attributed to measurement error. This operationalisation is favoured by Jacobson and colleagues [1999; Jacobson and Truax, 1991] and in about one-third of all studies of clinical significance, and is therefore the most often used [Ogles et al., 2001].

Only one study by Nixon et al. [2003] investigated in this manner the clinical significance of intervention effects in children with externalizing behaviour disorders. However, this study looked at clinically disturbed children, and it involved a therapy, not a prevention measure, evaluated in a small sample ($n = 16$).

In the above-mentioned application study of PEP effects under routine care conditions, the clinical significance of these effects was verified. Whereas before the PEP parent training, depending on the assessment instrument, 32.6–60.7% of the children displayed clinically significant disorders, and 3 months later, depending on the assessment instrument, 24.8–60.4% of these children were considered clearly improved [Hautmann et al. 2009b]. Such clinically significant changes are, from the standpoint of preventing mental disorders [Cuijpers et al., 2005], especially of practical relevance if they persist in the longer term.

The aim of the present study is to verify whether, 2 years after participating in a PEP training course, there are still clinically significant changes in children's problem behaviour and their parents' parenting behaviour. It is assumed that clinical significance is also an appropriate measure to assess the practical usefulness of preventive programmes. Based on the positive short-term effects of the training programme, we also believe that long-term effects can be demonstrated.

Method

Description of Sample

Out of a sample of 2,121 Cologne preschool children, 243 families were identified based on the CBCL 4–18, a 13-item comprehensive screening tool [Plück et al., 2008]; the 3- to 6-year-old children, in the assessment of their parents and teachers, displayed externalizing problem behaviour (indicated sample; fig. 1). Of these 243 families, 88 refused further partici-

pation or could not participate further because of language difficulties. The remaining 155 families were divided into 90 for the intervention group and 65 for the control group. Randomisation was done by kindergarten. Of the 90 families, 30 did not receive parent training; that group received only teacher training. The intervention and control groups did not differ in socio-demographic data, severity of the child's problem behaviour or parenting skills [Hanisch et al., 2006]. For the present analysis, longitudinal data were available from 34 control group families (CG), 38 families in the group in which both parents and teachers received PEP training (P+T) and 21 children whose teachers completed the PEP training (T). Table 1 describes the sample.

Study Design

The study was designed as a randomised controlled trial, for which data were collected for all groups before the start of training (pre), after the end of training, and a maximum of 3 months after pre-testing in the CG (post), 6 months after PEP training (follow-up 1), 18 months after training (FU2) and 30 months after training (FU3). The CG received no intervention.

The pre and post assessments comprised a 2- to 3-hour home visit and questionnaire booklets for mother, father and teacher. For the follow-up measurements, parents and preschool teachers were again given a ques-

tionnaire, or schoolteachers were given it if the child had already begun school. Written consent for participation in the study was obtained from all subjects. The study was approved by the Ethics Committee of the University Hospital of Cologne.

Intervention

PEP is primarily intended for parents and teachers of 3- to 6-year-old children with externalizing problem behaviour or clinically relevant externalizing behaviour disorders. There is a training manual for PEP [Plück et al., 2006]. It consists of ten 90- to 120-minute weekly sessions, each attended by 5-6 people, in separate groups for parents and teachers. In the first 3 sessions, individual problem situations are defined and strategies are conveyed for how parents and teachers can strengthen their relationship with the child. The model shows how operant mechanisms contribute to the perpetuation of problem behaviour, in a vicious cycle. Sessions 4-6 give the parents and teachers the basic strategies of behaviour modification, using predefined individual problem situations: laying out the rules, formulating requests, positive consequences for desired behaviour and negative consequences for problem behaviour. The remaining 4 sessions are used to apply the strategies learned to other difficult parenting or teaching situations. The PEP attaches great importance to being easily understandable and practical in everyday life, so as to be able to reach poorly educated parents.

Rating Instruments

Long-term effects of PEP training on parenting behaviour and on the child's problem behaviour are subsequently described from the mother's point of view.

Positive parenting, from the mother's point of view: The questionnaire *Fragen zum Erziehungsverhalten* (FZEV) is a German adaptation and further elaboration of the Parent Practices Scale [Strayhorn and Weidmann, 1988] and was developed by a group at the University of Brunswick, in co-operation with us. The questionnaire contains 13 items that measure positive, reinforcing and supportive parenting on a 4-point scale. In our sample the internal consistency of the general scale was $\alpha = 0.84$.

The questionnaire *Erziehungsfragebogen* (EFB) [Miller, 2001] (English version, The Parenting Scale; Arnold et al., 1993) includes 29 items that deal with dysfunctional parenting strategies. Parents assess their tendency to use certain parenting strategies on a 7-point scale. The internal consistency for the total value in our sample was $\alpha = 0.76$.

The questionnaire *Verhalten in Risikosituationen* (VER) [Hahlweg and team, Brunswick], the German adaptation of The Triple P-Positive Parenting Program [Sanders et al., 2000], asks about the subjective ability to cope with difficult parenting situations. Parents are supposed to judge

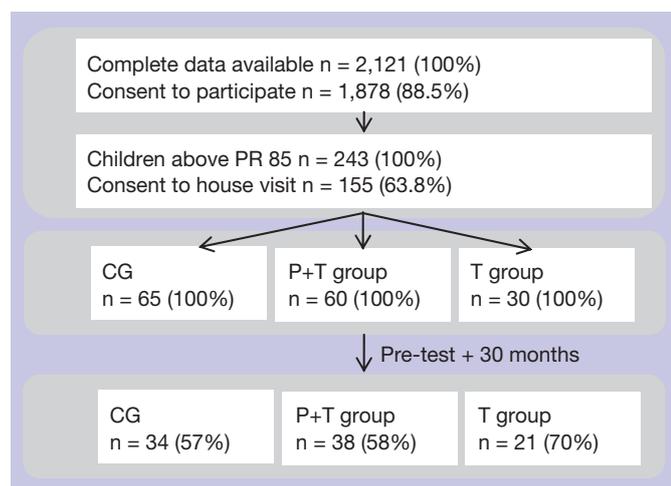


Fig. 1. Recruitment of the indicated sample; PR = percentile rank.

Table 1. Description of the sample 30 months after PEP training, as well as statistical parameters to compare the PEP parent and teacher training group (P+T), the control group (CG) and the PEP teacher training group (T), before the start of training

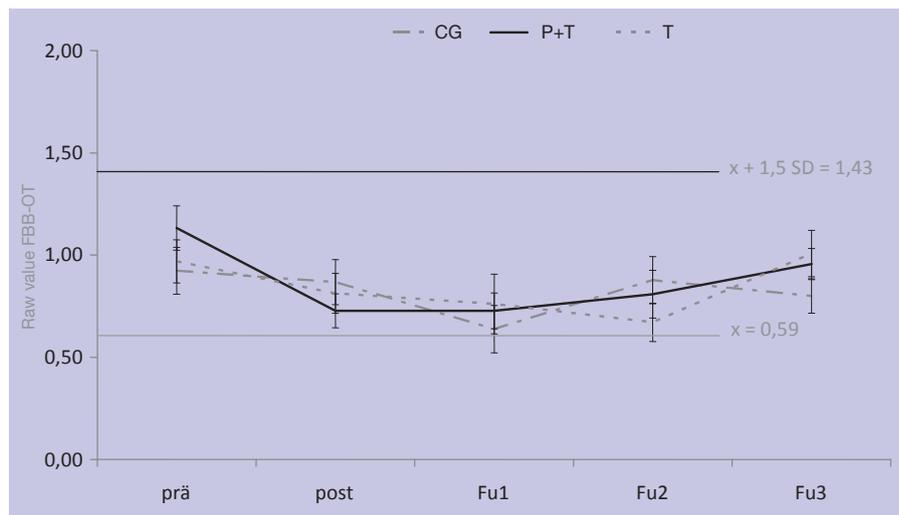
	P+T n = 38	CG n = 34	T n = 21	F	p
Age of child, years (mean)	4.11	4.21	4.38	0.69	0.51
Education	-0.23	0.20	-0.06	2.55	0.08
Symptoms	0.13	-0.13	-0.10	0.88	0.42
				Chi ²	p
<i>School education^a</i>					
Mother	2.35	2.03	1.59	5.54	0.063
Father	2.12	1.90	1.55	4.86	0.089
<i>Vocational training^b</i>					
Mother	1.03	0.78	0.56	5.65	0.059
Father	1.12	1.03	0.72	3.61	0.17

^a0 = No graduation, 1 = secondary general school,

2 = intermediate secondary school, 3 = academic high school or higher.

^b0 = Unskilled/semi-skilled, 1 = completed vocational training, 2 = college degree.

Fig. 2. Group mean values and standard error of the FBB-OT at the 5 measurement points. Shown are the group mean values used for normalisation $x = 0.59$ and $x + 1.5 s$ ($s = 0.56$), which is often used as a value for determination of clinical abnormality [Döpfner et al., 2008]. P+T = Parent and teacher training group, CG = control group, T = teacher training group, pre = pre-test, post = test after training, FU1 = 6 months after training, FU2 = 18 months after training, FU3 = 30 months after training.



whether they are able to cope with 27 difficult situations on a scale from 'not at all' (= 1) to 'very well' (= 4). The internal consistency in our sample was $\alpha = 0.90$.

The questionnaire *Fragen zur Selbstwirksamkeit* (FSW) [Hahlweg and team, Brunswick] is the German adaptation of the Parenting Sense of Competence Scale [Johnston and Mash, 1989]. The questionnaire contains 15 items on parental self-efficacy to be assessed on a 4-point scale. In this sample, the internal consistency was measured at $\alpha = 0.80$.

Problem behaviour of the child from the parents' point of view: The questionnaire *Fremdbeurteilungsbogen für Aufmerksamkeitsdefizit-/Hyperaktivitätsstörungen* (FBB-ADHS) consists of 23 questions to rate attention deficit hyperactivity disorders in children, and the *Fremdbeurteilungsbogen für Störung des Sozialverhaltens* (FBB-SSV) consists of 24 questions to rate conduct disorder in children; these measures satisfy the DSM-IV and ICD-10 diagnostic criteria for hyperkinetic disorder/attention deficit hyperactivity disorder or conduct disorder [Döpfner et al., 2008b]. In this study, only the subscale *Oppositionelles Trotzverhalten* (FBB-OT), which measures oppositional defiant disorder, was used; it consists of 9 items. The internal consistency of the mothers' assessments in this sample was $\alpha = 0.91$ for the total value on the questionnaire *Fremdbeurteilungsbogen für hyperkinetische Störungen* (FBB-HKS), which measures hyperkinetic disorders, and $\alpha = 0.88$ on the FBB-OT scale.

The questionnaire *Elternfragebogen über das Verhalten von Klein- und Vorschulkindern* (CBCL 1½-5; Arbeitsgruppe Deutsche Child Behavior Checklist, [2002]) is the German version of the Child Behavior Checklist for Ages 1½-5 [Achenbach and Rescorla, 2000]. The form contains 99 items that are scored from 0 (= not applicable) to 2 (= accurate or often apply). The items cover a wide range of behavioural disorders. In our sample, the internal consistency for the general scale was $\alpha = 0.94$. To consider changes in the individual raw scores, the CBCL 1½-5 was used throughout the project period. A comparison with existing standard values was no longer possible 30 months after training (FU3), due to the age of the children.

To reduce the number of tests by reducing the large number of dependent variables, the general scales of the questionnaires for mothers were combined, using factor analysis, into two general scales [Hanisch et al., 2006]. The mean of the z-standardised scale values of the FZEV, EFB, VER and FSW was the general scale *Positive Parenting*. The general scale *Problem Behaviour of the Child* was formed from the z-standardised values of the FBB-DHS and FBB-OT and the CBCL 1½-5 scale of externalizing problem behaviour. The internal consistencies of these general scales were $\alpha = 0.64$ for Positive Parenting and $\alpha = 0.96$ for Problem Behaviour of the Child.

Statistical Analyses

Missing data were addressed in three ways: scale values were only calculated if less than 10% of the relevant items were missing. Missing pre-test scale values were replaced by the mean value of the respective group [Tabachnick and Fidell, 1996]. Measurement points after the pre-test were always replaced using regression analysis, if the next test after the missing measurement point was available.

Pre-test values of intervention groups and CG were compared by the Kruskal-Wallis test or variance analysis.

Long-term effects of training were tested by repeated measure analysis of variance (GLM). Paired t-tests were then calculated by group, to study intra-group changes in the general scales of Problem Behaviours of the Child and Positive Parenting from pre to FU3.

A change is considered to be clinically significant if a person formerly classified as clinically abnormal is classified as normal after treatment. A prerequisite for determining clinical significance is therefore initial clinical abnormality. In the indicated sample studied here, children were included who scored in the ≥ 85 th percentile rank (PR) in a screening test. Thus, the sample included clinically disturbed children and children with subclinical externalizing problem behaviour. The children in the group mean are above the normal mean on the FBB-ADHS and FBB-OT [Döpfner et al., 2008b], but well below clinical abnormality (normal mean value plus 2 standard deviations [s]) (fig. 2). Significant problem behaviour in the borderline clinical range [Working Group on the German Child Behavior Checklist, 2002] was defined here as a deviation of ≥ 1.5 standard deviations from the norm mean value. Table 2 shows how many children, at the 5 measurement points, showed an externalizing problem behaviour defined in this way.

In addition to this return to a normal level of functioning, the RCI was defined as [Jacobson et al., 1999; Jacobson and Truax, 1991]:

$$\frac{X_{t2} - X_{t1}}{\sqrt{2(s_{t1} \sqrt{1 - r_{xx}})^2}}$$

where X_{t2} is the individual raw score at time 2 and X_{t1} is the individual raw score at time 1. s_{t1} is the standard deviation of the sample at time 1, and r_{xx} is defined as the reliability of the assessment method. If the RCI > 1.96 , the change is considered significant at the 0.05 level. A child is considered improved to a clinically significant extent, if the RCI > 1.96 and the value for the FBB-ADHS or FBB-OT at time FU3 lies in the normal range.

Respecting clinical significance, the RCI was used separately to rule out coincidence in the measurement of individual improvement, and this change is considered independently of the starting value. For each scale of children's problem behaviour (CBCL, FBB-ADHS, FBB-OT), sepa-

Table 2. Percent of children who were ≥ 1.5 standard deviations above the normal mean value in pre or FU3 in the FBB-AADHS or FBB-OT. Also shown is the number (%) of children who were assessed as clinically significantly improved from pre to post or FU3, and the number (%) of children who deteriorated from pre to post or FU3 from a value of ≤ 1.5 standard deviations to a value of ≥ 1.5 standard deviations above the mean.

	Disturbed children, n (%)			Clinically signif. improvement, n (%)		Incidence, n (%)	
	Pre	Post	FU3	Pre/Post	Pre/FU3	Pre/Post	Pre/FU3
<i>CG (n = 34)</i>							
FBB-ADHS	11 (32.4)	10 (29.4)	8 (23.5)	7 (20.6)	11 (32.4)	5 (14.7)	5 (14.7)
FBB-OT	6 (17.5)	7 (20.6)	5 (14.7)	7 (20.6)	6 (17.6)	7 (20.6)	4 (11.8)
<i>P+T (n = 38)</i>							
FBB-ADHS	14 (36.8)	11 (28.9)	7 (18.3)	14 (36.8)	13 (34.2)	7 (18.4)	5 (13.2)
FBB-OT	15 (39.5)	3 (7.9)	3 (7.9)	15 (39.5)	13 (34.2)	1 (2.6)	2 (5.3)
<i>T (n = 21)</i>							
FBB-ADHS	7 (33.3)	3 (14.3)	2 (9.5)	9 (42.9)	6 (28.6)	2 (9.5)	2 (9.5)
FBB-OT	4 (19.0)	3 (14.3)	3 (14.3)	4 (19.0)	6 (28.6)	2 (9.5)	3 (14.3)

Table 3 z-Standardised mean value (s) of the total scale symptoms and education at all time points; paired t-test comparisons of pre/post and pre/FU3

Group	Pre	Post	FU1	FU2	FU3	t (Pre/Post)	p (Pre/Post)	t (Pre/FU3)	p (Pre/FU3)
<i>CG (n = 34)</i>									
Symptoms	-0.13 (0.91)	-0.26 (0.97)	-0.60 (0.76)	-0.40 (0.88)	-0.50 (0.65)	0.52	0.66	1.95	0.06
Education	0.19 (0.82)	0.47 (0.62)	0.26 (0.84)	0.43 (0.89)	0.40 (0.65)	-1.44	0.16	-1.36	0.18
<i>P+T (n = 38)</i>									
Symptoms	0.13 (0.93)	-0.53 (0.78)	-0.49 (0.75)	-0.53 (0.84)	-0.33 (0.62)	3.37	0.002*	2.46	0.02*
Education	-0.23 (0.82)	0.43 (0.62)	0.31 (0.56)	0.38 (0.81)	0.08 (0.59)	-3.81	0.001*	-1.90	0.07
<i>T (n = 21)</i>									
Symptoms	-0.10 (0.75)	-0.62 (0.61)	-0.51 (0.65)	-0.31 (0.57)	-0.06 (0.73)	2.53	0.02*	1.01	0.33
Education	-0.06 (0.73)	0.38 (0.67)	0.32 (0.93)	0.62 (0.65)	0.27 (0.70)	-1.83	0.08	-1.51	0.15

*p \leq 0.05.

rate RCI values were calculated first, and then averaged to one RCI. Then the subjects were classified according to their RCI as showing 'significant deterioration' ($RCI < -1.96$), 'no change' ($-1.96 < RCI < 1.96$) and 'significant improvement' ($RCI > 1.96$). Group differences were compared using the Kruskal-Wallis test.

Correlations between the severity of the problem behaviour before the intervention and the RCI were determined using Pearson correlation coefficients.

In addition to the number of children who had clearly improved, table 2 also provides incidence rates that include children who, analogous to the above definition of significant problem behaviour, changed from an initially normal to an abnormal value.

Results

Preliminary analyses: Comparing those who dropped out of the study and the families who participated up to FU3, no significant differences were found with regard to the child's age or the severity of the problem or parenting behaviour (age: $F_{1,153} = 0.012$, $p = 0.912$; symptoms: $F_{1,153} = 0.206$, $p = 0.651$; education: $F_{1,153} = 0.056$, $p = 0.841$). There were, however, significant group differences in parental education: both

the fathers and the mothers of the families who participated to the end of the study had more education (father's education: $\chi^2_{(1)} = 5.75$, $p < 0.016$; mother's education: $\chi^2_{(1)} = 8.80$, $p < 0.003$).

Within the sample that was still active up to FU3, comparison among the three groups showed no significant pre-test differences in parenting behaviour, child behaviour problems, age of child or parental education (see table 1).

Figure 2 is exemplary of the group mean value of the raw scores of the FBB-OT at the 5 measurement points. The trends over time show no clear descriptive differences among the groups. But in the repeated measure analysis of variance (GLM), there was a clear time effect over the 5 measurement points for symptoms ($F_{4,87} = 5.17$, $p < 0.001$) and education ($F_{4,87} = 4.93$, $p < 0.001$), but no group \times time interaction effect (symptoms: $F_{8,176} = 0.937$, $p = 0.488$; education: $F_{8,176} = 0.937$, $p = 0.810$). Accordingly, the effect sizes of the CG group compared with the P+T group for pre versus FU3 were -0.05 for total symptom value and 0.1 for education.

Within each group, the following changes were observed (table 3): significant improvements in the child's problem be-

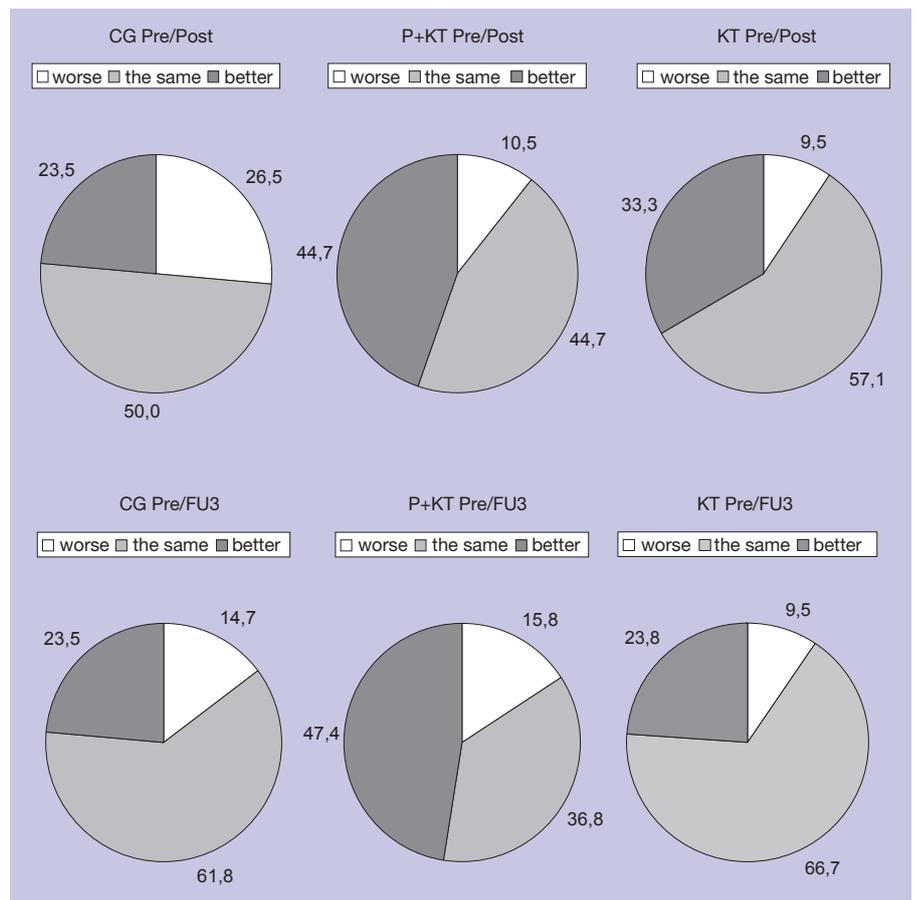


Fig. 3. Distributions of the Reliable Change Index (RCI) over the categories of improvement, deterioration and no change (in percent). The top row shows the change from pre to post; the bottom row shows the change from pre to FU3. P+T = Parent-and teacher training group, CG = control group, T = teacher training group.

behaviour were found between the pre and FU3 measurement in the P+T group, but not in the other two groups. No significant effect of parenting behaviour was found in any of the groups between the pre and the FU3 measurements. In the pre/post comparison, the P+T group showed significant improvements in both parenting behaviour and in the child's problem behaviour. In the T group, there was significantly reduced child problem behaviour at the post time point, but no change in parenting behaviour. In the CG, there were no significant changes from pre to post in either problem behaviour or parenting.

Table 2 shows that during the course of the intervention, more children showed clinically significant improvement in the intervention groups: 36.8% (ADHD) and 39.5% (OT) clinically significant improvements in the P+T group, 42.9% (ADHD) and 19% (OT) improvements in the T group, but only 20.6% in the CG. Over the entire span of the study (pre to FU3) the three groups became equalised: In the P+T group, 34.2% of the children displayed clinically significant improvements, in the T group 28.6% and in the CG 32.4% or 17.6%. The incidence rates initially suggested an advantage of the intervention groups compared with the CG group (pre/post), which, however, disappears between pre and FU3.

However, if we consider the RCI as a measure of reliable symptom reduction, independently of baseline values (see fig.

3), there is a clear improvement from pre to post of 44.7% in the P+T group, compared to 33.3% in the T group and 23.5% in the control group. Clear deterioration of 10.5% was found in the P+T group, 9.5% in the T group and 26.5% in the CG. Over the entire span from pre to FU3 in the P+T group, 47.4% of the children displayed a clear reduction of their problem behaviour, in the T group it was 23.8% and in the CG it was 23.5%. A comparison of the changes from pre to FU3, using the three-level factor ('significant improvement', 'no change' and 'significant deterioration'), with the help of the Kruskal-Wallis test, showed, however, no significant differences among the groups over time ($\chi^2_{(2)} = 2.83, p = 0.243$).

The correlation between the pre-test value for child's problem behaviour and the extent of change (RCI) showed a significant correlation of 0.772 ($p < 0.001$). Thus, the higher the initial level of problem behaviour, the stronger was the change from pre to FU3.

Discussion

The aim of this analysis was to verify whether the PEP can reduce externalizing problem behaviour in children over the long term and with clinically significant results. Considering all 5 measurement points, no group-specific trends in symp-

toms and parenting behaviour were found. At the level of the individual group, the group in which parents and the teacher participated in the PEP training, displayed clear short-term as well as long-term improvements in the child's problem behaviour. Parenting only improved over the short term in this group. Perhaps this result is due to a lack of consistency (0.64) on the general scale of Positive Parenting. In the group in which only the teachers were trained, and in the CG, as expected, there was no improvement in parenting. In the T group there were short-term improvements in the child's problem behaviour, but they did not last.

The positive effects of parental training programmes are well documented in statistical comparison of groups [Kaminski et al., 2008; Lundahl et al., 2006]. Verification of the clinical significance of these changes is less common [Nixon et al., 2003].

As a measure of clinically significant change, we used the combination of the RCI [Ogles et al., 2001] and return to a normal level of functioning [Jacobson et al., 1999, Jacobson and Truax, 1991]. More children in the P+T group showed clinically significant improvement from pre to post, while at 30 months after the training programme, a similar number of children in the CG were clearly improved. At the level of the RCI, the CG as well as the T group showed a clear reduction in externalizing behaviour, affecting nearly 24% of the children. Both cross-sectional and longitudinal studies support this finding, as they show greater prevalence rates of externalizing behaviour problems at the preschool and primary school age [Bongers et al., 2004]. Kuschel and colleagues [2004] report, on the basis of a German kindergarten sample of 3- to 6-year-olds, a 2–3 times higher prevalence of oppositional problem behaviour than in younger or older children. Longitudinal studies show that for a large portion of children who display subclinical externalizing problem behaviour at preschool age, the behaviour problems have decreased by early primary school age [Hartup, 2005; van Lier et al., 2007]. In the 5th measurement (FU3), the children of our sample were on average just under 7 years of age, and all were enrolled in school. Enrolment in school is a stressor, particularly for children with externalizing behaviour disorders [Webster-Stratton and Reid, 2006], that is likely to increase behavioural problems. In the P+T group, 47% of children show significant long-term improvement compared to the first test. Compared with the 24% of clearly improved children in the CG and taking into account the stressor of school enrolment, this can be evaluated as an improvement that has great practical relevance.

With respect to clinical significance, there were clear and similar improvements in the children's problem behaviour in both the intervention groups and in the CG. But it may be that PEP reduces the externalizing behaviour problems sooner, because the number of improved children is the same, in this case, from pre to post as from pre to FU3.

The only other study of clinically significant effects of intervention in children with externalizing problem behaviour

[Nixon et al., 2003] reported that 59% of the children improved significantly. Here, however, a parent-child interaction therapy was studied in a clinical sample, so the more severe symptoms and greater intensity of intervention had more pronounced effects.

To evaluate the effectiveness of a prevention programme, there should be both statistically and clinically significant long-term effects, both from a randomised controlled trial design and from application studies. Our prevention programme has proven effective in the short term in both study designs [Hanisch et al., 2010; Hautmann et al., 2008; 2009a], with the PEP intervention groups showing superiority in clinically significant improvements, especially over the short term. In the long term, stable intra-individual changes were documented in the combined intervention groups, using the RCI.

Long-term effects were only suggested, but we were unable to prove them at the level of statistical group comparison. Hence the need for urgent further research into the long-term effects of indicated prevention.

Many studies that deal with the long-term course of aggressive behaviour, describe the extremely unfavourable trend of so-called 'early starters'. 'Early starters' are characterised by an especially early, very pronounced pattern of externalizing behaviour, which also lasts through the elementary school years [van Lier et al., 2007; Hartup, 2005]. Early identification could be based either on known configurations of risk factors [Moffit, 1993] or on initial externalizing behaviour disorders [Plück et al., 2010]. Our data show a reduction of problem behaviour particularly among children whose symptoms were initially quite pronounced. In combination with the screening instrument, PEP could be a useful way to identify at-risk children and to strengthen their families' parenting skills. On the other hand, it may be that a specific training effect here coincides with the known phenomenon of regression toward the mean [Nesselroade et al., 1980] and the natural course of development of externalizing behaviour disorders [Bongers et al., 2004]. As evidence for a specific effect, however, a moderator analysis showed that the extent of the initial symptoms was the only predictive measure for short- and long-term therapeutic change in the effectiveness study [Hautmann et al., 2010].

It should be noted that a major limitation of our study is that it uses a selective sample to consider long-term effects. Right after the screening, some families refused to participate in the project [Hanisch et al., 2006], and we do not know to what extent this was a selective choice. Of the parents who participated in the project, we were able to keep only 62% in the project for nearly 3 years, and these tended to be the parents who were better educated than those who dropped out. Risk factor models of externalizing problem behaviour suggest that families with multiple psychosocial risk factors are over-represented in an indicated sample. The sample that was analysed up to the end of the project corresponds, at least in terms of secondary school qualifications and educational achievement, to the general population [Federal Statistical Office,

2009]. This suggests that the long-term sample represents a multiply selected subgroup of our initially selected sample. We can therefore draw no conclusions about the applicability of the results to those who dropped out of the study.

Another drawback, especially with respect to evaluating the clinical relevance of the results, is that the children were never clinically examined. Long-term data exist only for the mothers' judgement and the assessment by questionnaire.

It is also critical to look at the different starting levels of the three groups: Although the direct group comparison found no differences in the pre-test values of problem behaviour and parenting, the combined training group did show more significant child problem behaviour. These pre-test dif-

ferences also existed in the original sample, and therefore cannot be explained by a selective drop-out tendency of less severely impaired children in the CG or by selectively longer participation by more clearly stressed families.

Disclosure Statement

Manfred Döpfner and Julia Plück are authors of the PEP. Manfred Döpfner is director of the Ausbildungsinstitut für Kinder-Jugendlichenpsychotherapie (AKiP, Training Institute for Child and Adolescent Psychotherapy) at the University Hospital of Cologne, which offers training in PEP. Julia Plück conducts workshops in PEP. The other authors have no potential conflicts of interest.

References

- Achenbach TM, Rescorla LA: Manual for the ASEBA Preschool Form and Profiles. Burlington, University of Vermont, Department of Psychiatry, 2000.
- Arbeitsgruppe Deutsche Child Behavior Checklist: Elternfragebogen für Klein- und Vorschulkinder (CBCL 1½-5). Köln, Arbeitsgruppe Kinder-, Jugend- und Familiendiagnostik (KJFD), 2002.
- Arnold DS, O'Leary SG, Wolff LS, Acker MM: The Parenting Scale: a measure of dysfunctional parenting in discipline situations. *Psychol Assess* 1993;5: 137-144.
- Biederman J, Monuteaux MC, Mick E, Spencer T, Wilens TE, Silva JM, Snyder LE, Faraone SV: Young adult outcome of attention deficit hyperactivity disorder: a controlled 10-year follow-up study. *Psychol Med* 2006;36:167-79.
- Bongers I, Koot HM, van der Ende J, Verhulst FC: Developmental trajectories of externalizing behaviors in childhood and adolescence. *Child Dev* 2004; 75:1523-1537.
- Cuijpers P, van Straten A, Smit F: Preventing the incidence of new cases of mental disorders: A meta-analytic review. *J Nerv Ment Disord* 2005;193:119-125.
- Döpfner M, Breuer D, Schürmann S, Wolf Metternich T, Lehmkuhl G: Effectiveness of an adaptive multimodal treatment in children with attention-deficit hyperactivity disorder - global outcome. *Eur Child Adolesc Psychiatry* 2004;(suppl 1):117-129.
- Döpfner M, Schürmann S, Frölich J: Therapieprogramm für Kinder mit hyperkinetischem und oppositionellem Problemverhalten (THOP), ed 4. Weinheim, Beltz Psychologie Verlags Union, 2007.
- Döpfner M, Breuer D, Wille N, Erhart M, Ravens-Sieberer U; Bella Study group: How often do children meet ICD-10/DSM-IV criteria of attention deficit-/hyperactivity disorder and hyperkinetic disorder? Parent based prevalence rates in a national sample - results of the BELLA study. *Eur Child Adolesc Psychiatry* 2008a;17(suppl 1):59-70.
- Döpfner M, Görtz-Dorten A, Lehmkuhl G: Diagnostik-System für psychische Störungen im Kindes- und Jugendalter nach ICD-10 und DSM-IV (DISYPS-KJ II), ed 3. Bern, Huber, 2008b.
- Dretzke J, Frew E, Davenport C, Barlow J, Stewart-Brown S, Sandercock J, Bayliss S, Raftery J, Hyde C, Taylor R: The effectiveness and cost-effectiveness of parent training/education programmes for the treatment of conduct disorder, including oppositional defiant disorder, in children. *Health Technol Assess* 2005;9:iii, ix-x, 1-233.
- Elkins IJ, McGue M, Iacono WG: Prospective effects of attention deficit/hyperactivity disorder, conduct disorder, and sex on adolescent substance use and abuse. *Arch Gen Psychiatry* 2007;64:1145-1152.
- Eyberg SM, Nelson MM, Boggs SR: Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *J Clin Child Adolesc Psychol* 2008;37:215-237.
- Fontaine N, Carbonneau R, Barker ED, Vitaro F, Herbert M, Côté SM, et al: Girls' hyperactivity and physical aggression during childhood and adjustment problems in early adulthood: a 15-year longitudinal study. *Arc Gen Psychiatry* 2008;65: 320-328.
- Hanisch C, Plück J, Meyer N, Brix G, Freund-Braier I, Hautmann C, Döpfner M: Effekte des indizierten Präventionsprogramms für expansives Problemverhalten (PEP) auf das elterliche Erziehungsverhalten und auf das kindliche Problemverhalten. *Z Klin Psychol* 2006;35:117-125.
- Hanisch C, Freund-Braier I, Hautmann C, Jänen N, Plück J, Brix G, Eichelberger I, Döpfner M: Detecting effects of the indicated prevention programme for externalizing problem behaviour (PEP) on child symptoms, parenting, and parental quality of life in a randomized controlled trial. *Behav Cogn Psychother* 2010;38:95-112.
- Hartup WW: The development of aggression, in Tremblay RE, Hartup WW, Archer J (eds): *Developmental Origins of Aggression*. New York, Guilford, 2005.
- Hautmann C, Hanisch C, Mayer I, Plück J, Döpfner M: Effectiveness of the prevention program for externalizing problem behaviour (PEP) in children with symptoms of attention deficit/hyperactivity disorder and oppositional defiant disorder - generalisation to the real world. *J Neural Transm* 2008; 115:363-370.
- Hautmann C, Hoijtink H, Eichelberger I, Hanisch C, Plück J, Walter D, Döpfner M: One-year follow-up of a parent management training for children with externalizing behavior problems in the real world. *Behav Cogn Psychother* 2009a;29:379-396.
- Hautmann C, Stein P, Hanisch C, Eichelberger I, Plück J, Walter D, Döpfner M: Does parent management training for children with externalizing problem behavior in routine care result in clinically significant changes? *Psychother Res* 2009b;19:224-233.
- Hautmann C, Eichelberger I, Hanisch C, Plück J, Walter D, Döpfner M: The severely impaired do profit most: short-term and long-term predictors of therapeutic change for a parent management training under routine care conditions for children with externalizing problem behavior. *Eur Child Adolesc Psychiatry* 2010;19:419-430.
- Jacobson NS, Truax P: Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *J Consult Clin Psychol* 1991;59:12-19.
- Jacobson NS, Roberts LJ, Berns SB, McGlinchey JB: Methods for defining and determining the clinical significance of treatment effects: description, application, and alternatives. *J Consult Clin Psychol* 1999; 67:300-307.
- Johnston C, Mash EJ: A measure of parenting satisfaction and efficacy. *J Clin Child Psychol* 1989;18:167-175.
- Kaminski JW, Valle LA, Filene JH, Boyle CL: A meta-analytic review of components associated with parent training program effectiveness. *J Abnorm Child Psychol* 2008;36:567-589.
- Kirk RE: Promoting good statistical practices: some suggestions. *J Educ Psychol Meas* 2001;61:213-218.
- Kuschel A, Lübke A, Köppe E, Miller Y, Hahlweg K, Sanders MR: Häufigkeit psychischer Auffälligkeiten und Begleitsymptome bei drei- bis sechsjährigen Kindern: Ergebnisse der Braunschweiger Kindergartenstudie. *Z Kinder Jugendpsychiatr Psychother* 2004;32:97-106.
- Lundahl B, Risser HJ, Lovejoy MC: A meta-analysis of parent training: moderators and follow-up effects. *Clin Psychol Rev* 2006;26:86-104.
- Miller Y: Erziehung von Kindern im Kindergartenalter: Erziehungsverhalten und Kompetenzüberzeugungen von Eltern und der Zusammenhang zu kindlichen Verhaltensstörungen. Braunschweig, TU Braunschweig, 2001.
- Moffitt TE: Adolescent-limited and life-course-persistent antisocial behavior: a developmental taxonomy. *Psychol Rev* 1993;100:674-701.
- Mrug S, Windle M: Mediators of neighborhood influences on externalizing behavior in preadolescent children. *J Abnorm Child Psychol* 2009;37:265-80.
- Nelson G, Westhues A, MacLeod J: A meta-analysis of longitudinal research on preschool prevention programs for children. *Prev Treat* 2003;6:31.
- Nesselroade JR, Stigler SM, Baltes PB: Regression toward the mean and the study of change. *Psychol Bull* 1980;88:622-637.

- Nixon RDV, Sweeney L, Erickson DB, Touyz SW: Parent-child interaction therapy: a comparison of standard and abbreviated treatments for oppositional defiant preschoolers. *J Consult Clin Psychol* 2003;71:251–260.
- Ogles BM, Lunnen KM, Bonesteel K: Clinical significance: history, application, and current practice. *Clin Psychol Rev* 2001;21:421–446.
- Patterson GR, DeGarmo D, Forgatch MS: Systematic change in families following prevention trials. *J Abnorm Child Psychol* 2004;32:621–633.
- Plück J, Wieczorrek E, Wolff Metternich T, Döpfner M: Präventionsprogramm für Expansives Problemverhalten (PEP): Ein Manual für Eltern- und Erziehergruppen. Göttingen, Hogrefe, 2006.
- Plück J, Hautmann C, Brix G, Freund-Braier I, Hahlweg K, Döpfner M: Screening von expansivem Problemverhalten bei Kindern im Kindergartenalter für Eltern und Erzieherinnen (PEP-Screen). *Diagnostika* 2008;54:138–149.
- Plück J, Freund-Braier I, Hautmann C, Brix G, Wieczorrek E, Döpfner M: Recruitment in an indicated prevention program for externalizing behavior – parental participation decisions. *Child Adolesc Psychiatr Ment Health* 2010; accepted for publication.
- Sanders MR, Markie-Dadds C, Tully LA, Bor W: The Triple P-Positive parenting program: a comparison of enhanced, standard, and self-directed behavioral family intervention for parents of children with early onset conduct problems. *J Consult Clin Psychol* 2000;68:624–640.
- Statistisches Bundesamt (Hrsg): Statistisches Jahrbuch. Wiesbaden, Statistisches Bundesamt, 2009.
- Strayhorn JM, Weidmann CS: A parent practices scale and its relation to parent and child mental health. *J Am Acad Child Adolesc Psychiatry* 1988;27:613–618.
- Tabachnick BG, Fidell LF: Using multivariate statistics, ed 3. New York, Harper Collins, 1996.
- Tremblay RE: Prevention of youth violence: Why not start at the beginning? *J Abnorm Child Psychol* 2006;34:481–487.
- van Lier PA, van der Ende J, Koot HM, Verhulst FC: Which better predicts conduct problems? The relationship of trajectories of conduct problems with ODD and ADHD symptoms from childhood into adolescence. *J Child Psychol Psychiatr* 2007;48:601–608.
- Webster-Stratton C, Reid MJ: Treatment and prevention of conduct problems: parent training intervention for young children (2–7 years old); in McCartney K, Phillips DA (eds): *Blackwell Handbook on Early Childhood Development*. Malden, Blackwell, 2006, pp 616–641.