

Preventing Child Behavior Problems in the Erlangen-Nuremberg Development and Prevention Study: Results from Preschool to Secondary School Age

Friedrich Lösel, Institute of Criminology, Cambridge University, United Kingdom, and Institute of Psychology,
University of Erlangen-Nuremberg, Germany

Mark Stemmler, Institute of Psychology, University of Erlangen-Nuremberg, Germany

Vol. 6 (2) 2012

Evidence-based Developmental Prevention of Bullying and Violence in Europe

Editorial (p. 165)

Guest Editorial: The Future of Research on Evidence-based Developmental Violence Prevention in Europe – Introduction to the Focus Section Manuel Eisner / Tina Malti (pp. 166 – 175)

Violence Prevention in Austrian Schools: Implementation and Evaluation of a National Strategy Christiane Spiel / Petra Wagner / Dagmar Strohmeier (pp. 176 – 186)

Clinical Significance of Parent Training for Children with Conduct Problems Martin Forster / Åsa Kling / Knut Sundell (pp. 187 – 200)

From Clinical-Developmental Theory to Assessment: The Holistic Student Assessment Tool Gil Noam / Tina Malti / Martin Guhn (pp. 201 – 213)

► **Preventing Child Behavior Problems at Preschool Age: The Erlangen-Nuremberg Development and Prevention Study** Friedrich Lösel / Mark Stemmler (pp. 214 – 224)

Introducing, Researching, and Disseminating the Incredible Years Programmes in Wales Judy Hutchings (pp. 225 – 233)

Implementation of PATHS Through Dutch Municipal Health Services: A Quasi-Experiment Ferry X. Goossens / Evelien M. J. C. Gooren / Bram Orobio de Castro / Kees W. van Overveld / Goof J. Buijs / Karin Monshouwer / Simone A. Onrust / Theo G. W. M. Paulussen (pp. 234 – 248)

Effectiveness of a Universal School-Based Social Competence Program: The Role of Child Characteristics and Economic Factors Tina Malti / Denis Ribeaud / Manuel Eisner (pp. 249 – 259)

The Impact of Three Evidence-Based Programmes Delivered in Public Systems in Birmingham, UK Michael Little / Vashti Berry / Louise Morpeth / Sarah Blower / Nick Axford / Rod Taylor / Tracey Bywater / Minna Lehtonen / Kate Tobin (pp. 260 – 272)

Successful Bullying Prevention Programs: Influence of Research Design, Implementation Features, and Program Components Bryanna Hahn Fox / David P. Farrington / Maria M. Ttofi (pp. 273 – 282)

Tackling Cyberbullying: Review of Empirical Evidence Regarding Successful Responses by Students, Parents, and Schools Sonja Perren / Lucie Corcoran / Helen Cowie / Francine Dehue / D'Jamila Garcia / Conor Mc Guckin / Anna Sevcikova / Panayiota Tsatsou / Trijntje Völlink (pp. 283 – 292)

KiVa Antibullying Program: Overview of Evaluation Studies Based on a Randomized Controlled Trial and National Rollout in Finland Christina Salmivalli / Elisa Poskiparta (pp. 293 – 301)

Knowing, Building and Living Together on Internet and Social Networks: The ConRed Cyberbullying Prevention Program Rosario Ortega-Ruiz / Rosario Del Rey / José A. Casas (pp. 302 – 312)

Empowering Students Against Bullying and Cyberbullying: Evaluation of an Italian Peer-led Model Ersilia Menesini / Annalaura Nocentini / Benedetta Emanuela Palladino (pp. 313 – 320)

Identity Centrality and In-Group Superiority Differentially Predict Reactions to Historical Victimization and Harm Doing Rezarta Bilali (pp. 321 – 337)

A Farewell to Innocence? African Youth and Violence in the Twenty-First Century Charles Ugochukwu Ukeje / Akin Iwilade (pp. 338 – 350)

Open Section



Preventing Child Behavior Problems in the Erlangen-Nuremberg Development and Prevention Study: Results from Preschool to Secondary School Age

Friedrich Lösel, Institute of Criminology, Cambridge University, United Kingdom, and Institute of Psychology, University of Erlangen-Nuremberg, Germany

Mark Stemmler, Institute of Psychology, University of Erlangen-Nuremberg, Germany

A brief overview of the prevention part of the long-term Erlangen-Nuremberg Development and Prevention Study, which combines a prospective longitudinal and experimental design. Findings up to five years after intervention are reported. From a sample of 609 families with kindergarten children, subgroups participated in the universal prevention program EFFEKT (child social skills training, a parent training and a combination of both) or were assigned to equivalent control groups. The short-term evaluation showed significant effects in mediating constructs (social problem solving and parenting behavior) and in educators' ratings of children's social behavior. In a follow-up after two to three years, school report cards showed fewer children with multiple behavior problems. In a further follow up after four to five years program children reported fewer externalizing and internalizing problems than the control group. There were no significant effects in the mothers' reports on their children's behavior. Most significant effect sizes ranged between $d = 0.20$ and $d = 0.40$. The findings suggest various positive long-term effects of the intervention. However, one needs to be cautious with regard to over-generalizing the positive findings, because effect sizes vary over time and the positive findings could not be replicated in all investigated variables.

Introduction

In recent years development-oriented prevention of delinquency and violence has become a key topic of criminology and crime policy (Beelmann 2012; Farrington and Welsh 2007; Lösel and Bender 2012). Numerous programs have been created and implemented in families, schools, preschools, social services, clinics, and neighborhoods. The child age may range from pre-birth to adolescence and the prevention can be universal (for all members of a population), selected (for at-risk groups), or indicated (for children with pre-existing behavior problems). Many programs have broader targets such as preventing not only crime but also internalizing problems, substance misuse, and school dropout, or generally supporting a desirable child development.

The expansion of developmental prevention in criminology and related disciplines has sound reasons such as sub-

stantial prevalence rates of behavioral problems, problem stability in a small group of "early starters," frequent comorbidity of various disorders, difficulties of later treatment, and high monetary costs for society (Lösel 2012a). Early prevention programs are based on a range of theoretical concepts. For example, child social skills training programs address social information processing and problem-solving, which play an important role in aggressive behavior (e.g. Dodge and Pettit 2003). Parenting programs aim to reduce coercive interactions, inconsistency, and corporal punishment (e.g. Dishion and Patterson 1994). Parent- and child-oriented programs are based on social learning and parenting theories (e.g. Webster-Stratton, Reid, and Hammond 2004). Early home-visiting programs use theories of social learning and attachment (e.g. Olds et al. 2007). Other programs integrate various concepts in a systems- and development-oriented perspective (e.g. Haw-

The research reported in this article was funded by the German Federal Ministry of Family Affairs, Seniors, Women, and Youth; the German Red Cross;

the European Union; and other institutions. We wish to thank Andreas Beelmann, Sabine Beer, Doris Bender, Stefanie Hacker, Stefanie Jaurisch, Antje

Klindworth-Mohr, Daniela Runkel, and Susanne Wallner for their work on the project.

kins et al. 2008). Although the respective theoretical constructs form a sound basis for prevention programs, their empirical correlations with antisocial outcomes are mainly small to moderate (Hawkins et al. 1998), particularly with regard to protective effects (Lösel and Farrington 2012). Therefore, it would be unrealistic to expect very large effects even for sound developmental programs.

Overall, systematic reviews and meta-analyses have shown positive results (e.g. Beelmann 2012; Farrington and Welsh 2007; Lösel 2012a) and also desirable cost-benefit ratios (e.g. Aos et al. 2004). However, there is great heterogeneity in outcomes and the field is confronted with various problems: 1. Most programs are not based on empirical evaluation studies (e.g. Junger et al. 2007). 2. Even when randomized controlled trials or sound quasi-experimental evaluations are carried out, most studies have only short follow-up periods, so there is no information on the impact on youth crime (Farrington and Welsh 2012; Lösel and Beelmann 2003). 3. There are frequently difficulties reaching high-risk families, early dropout, and other implementation problems (Eisner and Meidert 2011; Lösel 2012a). 5. There is a particular lack of long-term systematic research outside the English-speaking world (Beelmann 2012; Lösel and Bender 2012).

Against this background we carried out the Erlangen-Nuremberg Development and Prevention Study (ENDPS). This project is one of the few examples that follow the recommendation of Farrington, Ohlin and Wilson (1996) to combine a prospective longitudinal design with an experimental program evaluation. The ENDPS started in 1999 and is to our knowledge currently the longest-running study with such a combined design in Europe. As requested by the editors of this special issue, the present article contains a brief overview of the evaluation part of the project. While we refer mainly to previously published empirical articles, hitherto unpublished findings will also be presented.

1. Method

1.1. Sample and Participation Rates

The ENDPS comprises a core study on universal prevention programs and various smaller evaluations of program adaptations for specific risk groups. The sample of

the core study consisted of 675 kindergarten children (336 boys, 339 girls) from 609 families at 61 kindergartens in the cities of Erlangen and Nuremberg in Bavaria, Germany. The average age of the children at the first measurement was $M = 4.7$ years ($SD = 9.3$ months). According to an index of socio-economic status (cf. Geißler 1994) the sample was very similar to the population of the area (Beelmann et al. 2006). In total, seven waves of data collection have been carried out to date (the fourth only in a small subsample). The first three waves took place annually, the others at longer intervals. The retention rates varied over time and with regard to type of data gathering (e.g., number of mothers' versus fathers' questionnaires). As usual, not all measures were completed by all participants. The attrition rate of families after the first three annual waves was 5 percent. In the sixth assessment (four to five years after the intervention) 85 percent of the families were retained. In the most recent seventh wave (nine to eleven years after the first one) the retention rate was about 90 percent. Because it sometimes took a long time to complete the assessment of all families, the later follow-up times have a bandwidth.

The sample for the program evaluation consisted of 282 children (age: $M = 4.6$ years, $SD = 8.8$ months). 9.4 percent of the families were lower class, 29.4 percent lower middle class, 42.0 percent middle class, 16.4 percent upper middle class, and 2.8 percent upper class (cf. Geißler 1994). Ten percent of all parents had foreign ethnicity and 11 percent were single mothers (there was one single father in the study). We grouped the children/families as follows: 1. child training; 2. parent training; 3. combination of child and parent training, and 4. control group. All training took place in the year after the first assessment and ended between two to three months before the second assessment. The assignment of families/children to the training groups and control groups followed both methodological and practical considerations. A random assignment on the individual level would have caused serious threats to validity, for example reactance of families not included in a program, experimental or compensatory rivalry of control group parents, and diffusion of treatment if training group and control group children/parents were in contact in the same kindergarten (Lösel 2007b). Furthermore, not all kin-

dergartens were suitable for the training with regard to distance to the families' homes, available space, and group size. To cope with such typical problems in prevention practice but achieve equivalence of training group and control group, our design followed several steps: First, the above-mentioned organizational criteria were used to select kindergartens suitable for training. Second, 21 training groups from training kindergartens were selected randomly from all suitable kindergarten groups. Third, matched pairs were recruited from the other kindergartens with regard to age, gender, socio-economic status, and pre-training behavioral problems as assessed by the educators (total score of the Social Behavior Questionnaire, SBQ; see instruments section). With regard to quantitative matching criteria we used the "untreated neighbor" with the closest score. Due to this procedure, pretest SBQ means in the training group and control group were clearly equivalent (0.02 *SD* difference; Lösel et al. 2009). To control other aspects for potential heterogeneity in group-wise randomization we applied not only ANCOVAs but also causal regression models that separate average, covariate, and conditional effects (Steyer et al. 2000).

The child skills program was offered to the parents of 190 children. Twelve children were not permitted to participate, leaving 178 children from 157 families who took part in the training. This represents a participation rate of 93.7 percent. Of the participating children 96 percent were present for at least half of the sessions. The parent training was offered to 255 families, of which 170 (67 percent) participated. In most cases the mothers represented the family ($n = 163$), but there were also 48 fathers in the courses (sometimes together with the mother). Three-quarters of the parents attended at least half of the program.

As there is no generally valid solution for dealing with dropouts in program evaluation, one component of our analyses focused on children and parents who had attended at least half of their program (Lösel et al. 2006). In addition, intent-to-treat analyses that allocated all dropouts to the TC were also carried out (Lösel et al. 2009). The respective results were rather similar. Although we observed no significant differences in matching variables between dropouts and completers, we nonetheless used equal n

comparisons in our evaluations; i.e. only individuals for whom the matched partner was available were entered in the analyses. This ensures equivalence between training group and control group.

1.2. Prevention Programs

After various pilot studies (e.g. Beelmann 2003) the following programs were chosen for the main evaluation in the ENDPS:

Child training: The training of children's social skills was based on "I Can Problem Solve" (Shure 1992), but updated and modified for the German context. It was delivered to twenty-one groups of six to ten children. The course is a manual-based group training in social problem-solving (Beelmann, Jaurusch, and Lösel 2004). The first part addresses verbal concepts, identification of emotions, and reflection on reasons for behavior. The second part contains training in problem-solving skills such as providing alternative solutions in conflicts, anticipation of actions and evaluation of consequences. The training uses a range of didactical methods. Each of the fifteen sessions lasted 30 to 60 minutes and there were three to five sessions per week (guided by two trained facilitators from the ENDPS).

Parent training: The parenting program was delivered in twelve courses in the afternoon or evening. Child-care was provided to enable parents to attend. The training aims to enhance positive parenting skills (Beelmann and Lösel 2004). It is partly based on the programs of the Oregon Social Learning Center (Dishion and Patterson 1996; Fisher et al. 1997) and was updated and adapted to the German context. Pilot studies suggested keeping the program short to increase participation and reduce dropout. The training consisted of five 90- to 120-minute sessions spread over five weeks. The courses were delivered by two experienced facilitators from our team (group sizes 6 to 15). The content included issues of positive parenting, requests and demands, setting limits, dealing with difficult parenting situations, coping with stress, and enhancing the family's social relationships. Structured presentations, group discussions, role-playing, self-awareness exercises, homework, and other didactic measures were used.

1.3. Instruments

The assessments of the children and families varied over time and according to the children's age (e.g., the problem-solving test for six-years olds contained more items than for three-years olds). They employed multiple methods and data sources, ranging from parent interviews and questionnaires through child assessments and kindergarten staff's ratings to school report cards and pediatric data (Lösel et al. 2005). As the focus of this article is on program evaluation, we mainly refer to the children's behavioral problems as reported by various informants. In addition, data on process evaluation and theoretically relevant mediating factors for program outcome are briefly reported (i.e. children's social problem-solving, and parenting behavior).

For *outcome evaluation* the children's behavioral problems were measured using our German adaptations (Lösel, Beelmann, and Stemmler 2002) of the Social Behavior Questionnaire (SBQ; Tremblay et al. 1992). We used the SBQ to gather independent information from kindergarten educators, mothers, and (when they were older) children's self-reports. Behavioral problems in the first and second grades of elementary school were assessed by a content analysis of the school report cards (Stemmler et al. 2005). Teachers' comments on behavioral and emotional aspects such as aggression, hyperactivity, emotional tone, and obeying rules were reliably categorized and used as indicators of behavioral problems. From the sixth wave onwards we also applied short versions of a German self-reported delinquent behavior scale (the Delinquenz-Belastungsskala, DBS; Lösel 1975).

Parenting behavior was measured by mothers' reports on our German adaptation of the Alabama Parenting Questionnaire (Shelton, Frick, and Wotton 1996). The children's social problem-solving competence was assessed via the German version of Spivak and Shure's (1982; Shure, 1990) Preschool Interpersonal Problem-Solving Test (PIPS; Döpfner, Lorch, and Reihl 1989). In this test the children are asked for interpretations, potential motives, and behavioral alternatives in response to conflict scenarios presented in pictures.

For *process evaluation* of the parent training we used anonymous ratings on five-point scales that addressed as-

pects such as the selection of topics, comprehensibility, usefulness for participants' own parenting, and overall user satisfaction (Lösel et al. 2005). The implementation of the child training was assessed via ratings of the children's on-task and off-task behavior in each session (Cangelosi 1996).

2. Results

Most of the following results are presented as effect sizes using Cohen's d coefficients. When the outcome measure had already been used in the pretest (e.g. the SBQ), we calculated the net difference between standardized effects in the training group and control group. When an outcome measure could only be measured at follow up (e.g. the DBS), d was based on the standardized difference at the respective measurement point.

2.1. Implementation Quality

The above-mentioned participation and completion rates suggest that the implementation of the programs was satisfactory. This was particularly the case for the child program, where both rates were clearly above 90 percent. More detailed process evaluations also indicate that the implementation was appropriate (Lösel et al. 2005). On average the participants of the parent training rated all aspects between "very good" (1) and "good" (2) with a variation from $M = 1.20$ to 1.74 . Overall satisfaction with the training was $M = 1.73$ ($SD = 0.43$), the quality of the facilitators was rated particularly positively ($M = 1.20$, $SD = 0.33$).

The children's behavior during the program was also satisfactory. Over all sessions the rates of relevant on-task behavior were between 75 and 80 percent ($M = 79.2$ percent, $SD = 16.7$). Disruptive off-task behavior was observed much less commonly ($M = 7.2$ percent, $SD = 8.3$). Only three children (2.9 percent) had on-task scores of less than 50 percent. These and other process data suggest that the programs were implemented at high quality. Therefore, a potential lack of effects in the outcome evaluation could not be attributed to poor implementation quality.

2.2. Effects on Mediating Factors

Our findings show significant program effects on theoretically mediating proximal factors. The child program had a positive influence on social problem-solving as measured

by the PIPS (Lösel and Beelmann 2005). Children in the training group produced a larger overall number of conflict solutions and a smaller proportion of aggressive solutions, and made fewer aggressive decisions ($d = 0.25-0.47$). Such effects were not observed when only the parents had participated in parent training. Although our project showed an overall improvement of social information processing and problem-solving with increased age (Beelmann, Lösel, and Stemmler 2010), the child social skills training seems to accelerate this process.

Similarly, there were specific effects of parent training on parenting behavior and attitudes (Stemmler et al. 2007). Shortly after the training, mothers from the training group reported significantly more positive parenting ($d = 0.30$) and less inconsistent discipline ($d = 0.29$) than the control group. The latter effect remained stable during the first year. There was no overspill of impact from mothers to fathers who did not participate in the program.

As expected, characteristics of the children's social information processing and of the parents' educational behavior were related to child behavior problems, although most correlations were small (Beelmann et al. 2010; Stemmler et al. 2007).

2.3. Short-Term Effects on Child Behavior (Two to Three Months)

The first outcome assessment of child behavior was carried out two to three months after the training. The kindergarten educators' ratings of the children showed a significant positive effect of the total EFPEKT program in the SBQ-total score ($d = 0.30$). The effects of the various program components were $d = 0.26$ for the child training, $d = 0.22$ for the parent training, and $d = 0.39$ for the combined parent and child training (Lösel, Beelmann, et al. 2006). These results remained consistent when we used causal regression models instead of covariance analyses (Lösel et al. 2009). The effects in the subscales on externalizing and internalizing problems were significant for the total program ($d = 0.17$ and 0.19), the child training (0.25 and 0.26), and the combined training (0.36 and 0.33), but not for the parent training (0.11 and 0.09). There were also some conditional effects showing that those children who had more behavioral problems before

the program benefitted most (Lösel et al. 2009). The effect sizes for this subgroup of those in greatest need ranged between $d = 0.25$ and 0.66 (partially due to a slight increase of problems in the respective control groups; Lösel, Beelmann et al. 2006).

In contrast to the educators' ratings, there was no outcome in the positive direction in the mothers' ratings of child behavior (Lösel et al., 2009). We even found a small negative effect ($d = -0.22$) on internalizing problems.

2.4. Long-Term Effects on Child Behavior (Two to Five Years)

Two years after the training we analyzed the content of the report cards at elementary school. This is a particularly valid outcome indicator because the teachers did not know who participated in the program at kindergarten age. Overall, there were significantly less behavioral problems in the training group than in the control group ($d = 0.17$; Lösel, Beelmann et al. 2006). The specific effects of the parent training and combined training were not significant, but the child skills training had a significant and substantial effect ($d = 0.35$). In further analyses we focused on those children for whom the teachers reported at least three behavioral problems at grade one and two (Lösel et al. 2009). Because these children showed relatively stable problems two to three years after the program they are at high risk for longer-term deviance. The results for the total training and for the combined training showed significant positive effects (1.2 percent vs. 4.3 percent and 0.0 percent vs. 6.1 percent in the training group vs. control group).

In contrast to this non-reactive information from the school teachers, the mothers' reports did not reveal any significant program effect (Lösel et al. 2009). The above-mentioned slightly negative short-term effect in the mothers' reports of internalizing child symptoms had also disappeared.

In the follow up four to five years after the program we were able to assess children's self-reports. Table 1 contains the findings on all trained children and their parents versus the control group in the SBQ scales and delinquency self-report (DBS).

Table 1: Results for children's self-reports of behavioral and emotional problems four to five years after the intervention (all trained EFFEKT children and their parents)

	Program	Control	Difference	Effect size
	<i>M</i> <i>SD</i>	<i>M</i> <i>SD</i>	<i>t-test</i>	<i>d</i>
SBQ - Total scale	9.26 (5.37)	11.29 (5.76)	3.07**	0.28
SBQ - Externalizing score	5.99 (5.37)	7.35 (4.99)	2.37*	0.28
SBQ - Internalizing score	3.26 (2.20)	3.94 (2.56)	2.57*	0.29
DBS - Delinquency scale	0.20 (0.47)	0.27 (0.67)	1.14	0.13

Notes:
d = Cohen's standardized mean difference
t-tests controlled for variance heterogeneity
 SBQ = Social Behavior Questionnaire
 DBS = Delinquency Self Report Scale
 total *n* = 282 (equal *n* in training group and control group)
 * *p* < 0.05; ** *p* < 0.01.

There were positive effects for the SBQ total problems score and the subscales on both externalizing and internalizing problems. The result in the DBS scale on self-reported delinquency went in the expected direction, but was not significant (perhaps due to the generally small numbers of offences at this age). The effects were similar for the subsamples rated above and below the SBQ median at Time 1. The outcomes for the different types of interventions varied. They were all positive for the combined parent and child training (*d* = 0.21–0.34), but due to the now smaller *n* they did not reach statistical significance. The results of the parent training in the SBQ were particularly positive and significant (*d* = 0.49–0.63), but not significant in the delinquency scale. The findings for the child training all went in the desirable direction (*d* = 0.19–0.45) and were significant in the DBS. The effects of the combined trainings in the delinquency scale were positive (*d* = 0.17–0.26), but not statistically significant.

We also recorded the mothers' evaluations of child behavior in the SBQ scales four to five years after the training. There

were no significant effects in the total score, nor in the subscales on externalizing and internalizing child problems.

3. Discussion

In comparison to the majority of evaluations on developmental prevention, follow-up periods of two to three and four to five years are rather long. Most studies comprise only one year or less (Beelmann 2012). This is particularly the case in Europe where well controlled long-term evaluations are very rare (e.g. Beelmann and Lösel 2007; Eisner et al. 2007).

A substantial proportion of our short- and long-term effects went in the desirable direction, i.e. less behavioral problems in the program groups. The significant effects were mainly small and occasionally moderate. This is in line with international meta-analyses that report lower effects for universal prevention than for risk-focused selective and indicated approaches (Beelmann and Raabe, 2009; Lösel 2012a). In principle, one should not expect larger long-term effects of universal programs because the majority of the children in the more or less "normal" intervention groups would not develop behavioral problems without receiving a program. This does not imply that universal programs for the whole population of a kindergarten, school, or neighborhood should not be carried out. Universal programs have the advantage of avoiding potential stigmatization, do not require risk assessment procedures and can more easily be implemented in routine practice (e.g. in schools). As in public health programs, universal approaches can also have a particular impact on those groups who are most in need or at highest risk (Coid 2003). This was confirmed in our short-term outcomes and in the school report cards two to three years after the training. After four to five years we did not find such a differential effect for those at highest risk in the children's self-reports. This may have been due to the overall small correlations between different informants and reduced stability over time (Lösel et al. 2005; see also below).

We conclude from these findings that a relatively short universal prevention program such as EFFEKT can reduce child behavior problems. This is in accordance with meta-analyses that found no strong relationship between pro-

gram intensity (dosage) and effect size (Beelmann 2008; Lösel and Beelmann, 2003). However, the universal approach should be seen as a “foot in the door” with regard to lasting influences in high-risk groups. Universal programs would become too expensive for the society if large numbers of low-risk children and families received high-dosage measures (Foster et al. 2008; Offord 2000).

Despite a number of theoretically plausible and practically desirable outcomes of the EFFEKT program one should be aware of various issues. Firstly there was some inconsistency of effect sizes over time. Although, as expected, a number of effects decreased in the follow-ups, some outcomes were stronger in the long-term than in the short-term evaluation. There was also partial inconsistency of effects in different outcome measures. For example, the positive effects in the kindergarten educators’ SBQ reports, elementary school report cards, and children’s self reports clearly differed from the results in the mothers’ reports where we found no positive effects. Similarly, we observed certain inconsistencies in the outcomes of the different parts of the program. For example, the child training exhibited the largest effect after two to three years, whereas the parent training seemed to be more effective in the follow up after four to five years.

Some of these variations may be due to random fluctuation. Others may be due to numerous program, individual, contextual, and methodological factors (Lösel 2012a). One must be aware that programs are part of the child’s “natural” development in which (causal) risk and protective factors vary over time, accumulate, and interact with each other (Dodge et al. 2008; Lösel and Bender 2003). This can lead to complex patterns of influences. For example, the small negative effects in the short-term evaluation of the mother’s reports on internalizing child problems may have been due to a temporal increase of the mother’s sensitivity. Larger effect sizes in the follow-up could be due to “sleeping effects” that may result from greater experience in the application of training content. Certain inconsistencies between the data from the educators/teachers and from the mothers reflect the generally small correlations between different informants on child behavior problems (Achenbach 2006) that was also observed in the ENDPS (Lösel et al. 2005). The reports of the pro-

fessionals may be more reliable because they contain comparisons between children (which is often lacking in small families). Mothers’ and teachers’ reports can be influenced by a relatively stable general image of the child. As Lösel reports (2002), the longitudinal correlations between behavior ratings by the same informants were larger than the cross-sectional correlations between different informants. In another study of the ENDPS we only found significant effects of a child skills training when the teachers who rated the child were not the same before and after the program (Hacker et al. 2007).

Variations and partial inconsistencies in findings across times, subprograms, and outcome measures are rather common in evaluations of developmental prevention programs (even in some of the best long-term studies). For example, the FAST Track project found substantial variation in different outcomes (Conduct Problems Prevention Research Group 2002, 2004, 2010); most recently there were positive effects in official indicators of offending but not in self-reports. The Seattle Development Project reported positive effects of “Communities That Care,” but with variations between parts of the program and different outcome measures (e.g. Hawkins et al. 1999, 2008). The Montreal Prevention Experiment had less effects in short-term than in long-term evaluations (Tremblay et al. 1995; Vitaro, Brendgen, and Tremblay 2001). These and other examples suggest that one should not expect too simple and uniform messages from rather complex and multimodal long-term evaluations of developmental prevention programs. One should also be aware of the risk of “fishing for significances” when many variables and measurement times are included.

Our evaluations of EFFEKT revealed not only various long-term effects, but also no lasting negative outcomes. As McCord (2003) has shown, some programs can harm in spite of best intentions. Having various positive and no harmful impact justified disseminating the EFFEKT program into routine practice. Meanwhile more than one thousand facilitators have been trained in all parts of Germany on a non-profit basis. We also learnt that the standard program needs to be enriched with modules for subpopulations with specific needs. Therefore, we devel-

oped adapted versions that addressed cultural differences in parenting (e.g. in Turkish families) and also used simpler language. Ethnic minority families in deprived neighborhoods benefitted from this intervention (Runkel 2009; Runkel et al., forthcoming). Another adaptation addressed emotionally burdened or depressed mothers (Kötter et al. 2010). This program was evaluated in clinic contexts and showed positive effects on mothers' parenting and children's behavior (Bühler et al. 2011; Stemmler et al., forthcoming). In principle, we recommend working with the core structure of evidence-based programs, but adding and evaluating modules for specific family needs where necessary. Such a more individualized approach is supported by a meta-analysis of family-oriented prevention programs in Germany (Lösel, Schmucker et al. 2006).

4. Conclusions

The findings and experiences from our primary evaluations and research syntheses within the ENDPS lead us

to the following conclusions: First, universal developmental prevention programs such as EFFEKT can have positive long-term effects. Second, it is realistic to expect mainly small effect sizes, particularly in routine practice (as opposed to demonstration projects). Third, because of the high costs of persistent criminality even small effects may well pay off if only a few cases become more resilient. Fourth, more well-controlled and replicated evaluations of the long-term outcomes of both universal and targeted prevention programs are needed, particularly outside North America. Fifth, there is a need for more research on the outreach and implementation of programs in routine practice. Sixth, specific evidence-based programs must be more closely integrated into the broad range of routine services in practice. And finally, developmental prevention on the individual and micro-social level should be accompanied by approaches designed to reduce risks at the macro-level (e.g. in order to avoid social segregation).

References

- Achenbach, Thomas. 2006. As Others See Us: Clinical and Research Implications of Cross-Informant Correlations for Psychopathology. *Current Directions in Psychological Science* 15:94–98.
- Aos, Steve, Roxanne Lieb, Jim Mayfield, Marna Miller and Annie Pennucci. 2004. *Benefits and Costs of Prevention and Early Intervention Programs for Children and Youth*. Olympia, WA: Washington State Institute of Public Policy.
- Beelmann, Andreas. 2003. Effektivität behavioraler Elternt Trainingsprogramme: Ergebnisse zweier Pilotstudien zur Prävention dissozialen Verhaltens. *Psychologie in Erziehung und Unterricht* 50:310–23.
- Beelmann, Andreas. 2008. *The Effect of Parent Training Programs in the Prevention and Treatment of Antisocial Behavior in Childhood and Adolescence*. Paper presented at the 18th European Conference of Psychology and Law, 2–5 July 2008, Maastricht, the Netherlands.
- Beelmann, Andreas. 2012. The Scientific Foundation of Prevention: The Status Quo and Future Challenges for Developmental Prevention. In *Antisocial Behavior and Crime: Contributions of Developmental and Evaluation Research to Prevention and Intervention*, ed. Thomas Bliesener, Andreas Beelmann, and Mark Stemmler, 137–63. New York: Hogrefe.
- Beelmann, Andreas, Stefanie Jaurisch, and Friedrich Lösel. 2004. *Ich kann Probleme lösen: Soziales Trainingsprogramm für Vorschulkinder*. Universität Erlangen-Nürnberg: Institut für Psychologie.
- Beelmann, Andreas, Stefanie Jaurisch, Friedrich Lösel, and Mark Stemmler. 2006. Frühe universelle Prävention von dissozialen Entwicklungsproblemen: Implementation und Wirksamkeit eines verhaltensorientierten Elternt Trainings. *Praxis der Rechtspsychologie* 16:120–43.
- Beelmann, Andreas, and Friedrich Lösel. 2004. *Elternt raining zur Förderung der Erziehungskompetenz*. Universität Erlangen-Nürnberg: Institut für Psychologie.
- Beelmann, Andreas, and Friedrich Lösel. 2007. Prävention von externalisierendem Problemverhalten. In *Prävention und Gesundheitsförderung*, vol. 3, Kinder und Jugendliche, ed. Bernd Röhrle, 557–95. Tübingen: Deutsche Gesellschaft für Verhaltenstherapie.
- Beelmann, Andreas, Friedrich Lösel, and Mark Stemmler. 2010. Die Entwicklung von sozialer Informationsverarbeitung und die Vorhersage physischer Aggression im Vorschulalter. *Psychologie in Erziehung und Unterricht* 57:119–31.
- Beelmann, Andreas, Friedrich Lösel, Mark Stemmler, and Stephanie Jaurisch. 2006. Beurteilung von sozialen Verhaltensproblemen und Erziehungsschwierigkeiten im Vorschulalter: Eine Untersuchung zur deutschen Adaptation des Eyberg Child Behavior Inventory (ECBI). *Diagnostica* 52:189–98.
- Beelmann, Andreas, and Thomas Raabe. 2009. The Effects of Preventing Antisocial Behavior and Crime in Childhood and Adolescence: Results and Implications of Research Reviews and Meta-analyses. *European Journal of Developmental Science* 3:260–81.
- Bender, Doris, and Friedrich Lösel. 2006. Working With Violent Children in German Youth Services: Results of a Survey. In *Children Who Commit Acts of Serious Interpersonal Violence: Messages for Practice*, ed. Ann Hagell and Renuka Jeyarajah-Dent, 167–85. London: Jessica Kingsley.
- Bühler, Anneke, Charlotte Kötter, Stefanie Jaurisch, and Friedrich Lösel. 2011. Prevention of Familial Transmission of Depression: EFFEKT-E, A Selective Program for Emotionally Burdened Families. *Journal of Public Health* 19:321–27.
- Cangelosi, James S. 1996. *Classroom Management Strategies: Gaining and Maintaining Students' Cooperation*. New York: Longman.
- Coid, Jeremy W. 2003. Formulating Strategies for the Primary Prevention of Adult Antisocial Behaviour: "High Risk" or "Population" Strategies? In *Early Prevention of Adult Antisocial Behaviour*, ed. David P. Farrington and Jeremy W. Coid, 32–78. Cambridge, UK: Cambridge University Press.
- Conduct Problems Prevention Research Group. 2002. Evaluation of the First 3 Years of the Fast Track Prevention Trial with Children at High Risk for Adolescent Conduct Problems. *Journal of Abnormal Child Psychology* 19:553–67.
- Conduct Problems Prevention Research Group. 2004. The Effects of the Fast Track Program on Serious Problem Outcomes at the End of Elementary School. *Journal of Clinical Child and Adolescent Psychology* 33:650–61.

- Conduct Problems Prevention Research Group. 2010. Fast Track Intervention Effects on Youth Arrests and Delinquency. *Journal of Experimental Criminology* 6:131–57.
- Dishion, Thomas J., and Gerald S. Patterson. 1996. *Preventive Parenting with Love, Encouragement and Limits: The Preschool Years*. Eugene, OR: Castalia.
- Dodge, Kenneth A., Mark T. Greenberg, Patrick S. Malone, and Conduct Problems Prevention Research Group. 2008. Testing an Idealized Dynamic Cascade Model of the Development of Serious Violence in Adolescence. *Child Development* 79:1907–27.
- Dodge, Kenneth A., and Gregory S. Pettit. 2003. A Biopsychosocial Model of the Development of Chronic Conduct Problems in Adolescence. *Developmental Psychology* 39:349–71.
- Döpfner, Manfred, Ruth Lorch, and Dieter Reihl. 1989. Soziale Informationsverarbeitung in Konfliktsituationen – eine empirische Studie an Vorschulkindern. *Zeitschrift für Pädagogische Psychologie* 3:239–48.
- Eisner, Manuel, and Ursula Meidert. 2011. Stages of Parental Engagement in a Universal Parent Training Program. *Journal of Primary Prevention* 32:83–93.
- Eisner, Manuel, Denis Ribeaud, Rahel Jünger, and Ursula Meidert. 2007. *Frühprävention von Gewalt und Aggression*. Zürich: Rüeegger.
- Farrington, David P., Lloyd E. Ohlin, and James Q. Wilson. 1986. *Understanding and Controlling Crime: Toward a New Research Strategy*. New York: Springer.
- Farrington, David P. and Brandon C. Welsh. 2007. *Saving Children From a Life of Crime*. Oxford, UK: Oxford University Press.
- Fisher, Philip A., Elizabeth Ramsay, Karla Antoine, Kate Kavanagh, Allen Winebarger, J. Mark Eddy, and John B. Reid. 1997. *Success in Parenting: A Curriculum for Parents with Challenging Children*. Eugene, OR: Oregon Social Learning Center.
- Foster, E. Michael, Ronald J. Prinz, Matthew R. Sanders, and Cheri J. Shapiro. 2008. The Costs of a Public Health Infrastructure for Delivering Parenting and Family Support. *Children and Youth Services Review* 30:493–501.
- Geißler, Rainer, ed. 1994. *Soziale Schichtung und Lebenschancen in Deutschland*. Stuttgart: Enke.
- Hacker, Stefanie, Friedrich Lösel, Mark Stemmler, Stefanie Jaurisch, Daniela Runkel, and Andreas Beelmann. 2007. Training im Problemlösen (TIP): Implementation und Evaluation eines sozial-kognitiven Kompetenztrainings für Kinder. *Heilpädagogische Forschung* 23:11–21.
- Hawkins, J. David, Eric C. Brown, Sabrina Oesterle, Michael W. Arthur, Robert D. Abbott, and Richard F. Catalano. 2008. Early Effects of Communities That Care on Targeted Risks and Initiation of Delinquent Behavior and Substance Use. *Journal of Adolescent Health* 43:15–22.
- Hawkins, J. David, Richard F. Catalano, Rick Kosterman, Robert D. Abbott, and Karl G. Hill. 1999. Preventing Adolescent Health-risk Behaviors by Strengthening Protection During Childhood. *Archives of Pediatric and Adolescent Medicine* 153:226–34.
- Junger, Marianne, Lynette Feder, Joy Clay, Sylvana M. Coté, David P. Farrington, Kate Freiberg, Vicente Garrido, et al. 2007. Preventing Violence in Seven Countries: Global Convergence in Policy. *European Journal on Criminal Policy and Research* 13:327–56.
- Kötter, Charlotte, Mark Stemmler, Anneke Bühler, Stefanie Jaurisch, and Friedrich Lösel. 2010. Mittelfristige Effekte des Präventionsprogramms EFFEKT-E für emotional belastete Mütter und ihre Kinder unter besonderer Berücksichtigung psychosozialer Risikofaktoren. *Zeitschrift für Gesundheitspsychologie* 19:122–33.
- Lösel, Friedrich. 1975. *Handlungskontrolle und Jugenddelinquenz: Theoretische Integration und empirische Prüfung*. Stuttgart: Enke.
- Lösel, Friedrich. 2002. Risk/Need Assessment and Prevention of Antisocial Development in Young People: Basic Issues from a Perspective of Cautionary Optimism. In *Multiproblem Violent Youth*, NATO SPS series, ed. Ray Corrado, Ron Roesch, Stephen D. Hart, and Josef Gierowski, 35–57. Amsterdam: IOS Press.
- Lösel, Friedrich. 2007a. It's Never Too Early and Never Too Late: Towards an Integrated Science of Developmental Intervention in Criminology. *Criminologist* 35 (2): 1–8.
- Lösel, Friedrich. 2007b. Doing Evaluation in Criminology: Balancing Scientific and Practical Demands. In *Doing Research on Crime and Justice*, 2d ed., ed. Roy D. King and Emma Wincup, 141–170. Oxford, UK: Oxford University Press.
- Lösel, Friedrich. 2012a. Entwicklungsbezogene Prävention von Gewalt und Kriminalität: Ansätze und Wirkungen. *Forensische Psychiatrie, Psychologie und Kriminologie* 6:71–84.
- Lösel, Friedrich. 2012b. Offender Treatment and Rehabilitation: What Works? In *The Oxford Handbook of Criminology*, 5th ed., ed. Mike Maguire, Rod Morgan, and Robert Reiner 986–1016. Oxford, UK: Oxford University Press.
- Lösel, Friedrich, and Andreas Beelmann. 2003. Effects of Child Skills Training in Preventing Antisocial Behavior: A Systematic Review of Randomized Evaluations. *Annals of the American Academy of Political and Social Science* 587:84–109.
- Lösel, Friedrich, and Andreas Beelmann. 2005. Social Problem-Solving Programme for Preventing Antisocial Behaviour in Children and Youth. In *Social Problem Solving and Offending: Evidence, Evaluation and Evolution*, ed. Mary McMurrin and James McGuire, 127–43. Chichester, UK: Wiley.
- Lösel, Friedrich, and Andreas Beelmann. 2006. Child Skills Training. In *Preventing Crime: What Works for Children, Offenders, Victims, and Places*, ed. Brandon C. Welsh and David P. Farrington, 33–54. Dordrecht, NL: Wadsworth.
- Lösel, Friedrich, Andreas Beelmann, Stefanie Jaurisch, Ute Koglin, and Mark Stemmler. 2005. Entwicklung und Prävention früher Probleme des Sozialverhaltens: Die Erlangen-Nürnberg Studie. In *Möglichkeiten der Gewaltprävention*, ed. Manfred Cierpka, 201–49. Göttingen: Vandenhoeck und Ruprecht.
- Lösel, Friedrich, Andreas Beelmann and Mark Stemmler. 2002. *Skalen zur Messung sozialen Problemverhaltens bei Vorschul- und Grundschulkindern*. Universität Erlangen-Nürnberg: Institut für Psychologie.
- Lösel, Friedrich, Andreas Beelmann, Mark Stemmler, and Stefanie Jaurisch. 2006. Prävention von Problemen des Sozialverhaltens im Vorschulalter: Evaluation des Eltern- und Kindertrainings EFFEKT. *Zeitschrift für Klinische Psychologie und Psychotherapie* 35:127–39.
- Lösel, Friedrich, and Doris Bender. 2003. Protective Factors and Resilience. In *Early Prevention of Adult Antisocial Behaviour*, ed. David P. Farrington and Jeremy W. Coid, 130–204. Cambridge: Cambridge University Press.
- Lösel, Friedrich, and Doris Bender. 2006. Risk Factors for Serious Juvenile Violence. In *Children Who Commit Acts of Serious Interpersonal Violence: Messages for Practice*, ed. Ann Hagell and Renuka Jeyarajah-Dent, 42–72. London: Jessica Kingsley.
- Lösel, Friedrich, and Doris Bender. 2012. Child Social Skills Training in the Prevention of Antisocial Development and Crime. In *Handbook of Crime Prevention*, ed. David P. Farrington and Brandon C. Welsh, 102–29. Oxford, UK: Oxford University Press.
- Lösel, Friedrich, and David P. Farrington. 2012. Direct Protective and Buffering Protective Factors in the Development of Youth Violence. *American Journal of Preventive Medicine* 43 (2, supplement): 8–23.
- Lösel, Friedrich, Martin Schmucker, Birgit Plankensteiner, and Maren Weiss. 2006. *Bestandsaufnahme und Evaluation der Elternbildung*. Berlin: Bundesministerium für Familie, Senioren, Frauen und Jugend.
- Lösel, Friedrich, Mark Stemmler, Andreas Beelmann, and Stefanie Jaurisch. 2005. Aggressives Verhalten im Vorschulalter: Eine Untersuchung zum Problem verschiedener Informanten. In *Aggressionsentwicklung zwischen Normalität und Pathologie*, ed. Inge Seiffge-Krenke, 141–67. Göttingen: Vandenhoeck und Ruprecht.
- Lösel, Friedrich, Mark Stemmler, Stefanie Jaurisch, and Andreas Beelmann. 2009. Universal Prevention of Antisocial Development: Short- and Long-Term Effects of a Child- and Parent-Oriented Program. *Monatsschrift für Kriminologie und Strafrechtsreform* 92: 289–308.
- McCord, Joan. 2003. Cures That Harm: Unanticipated Outcomes of Crime Prevention Programs. *Annals of the American Academy of Political and Social Science* 587:16–30.
- Offord, David R. 2000. Selection of Levels of Prevention. *Addictive Behaviors* 25:833–42.

- Olds, David L., Harriet Kitzman, Carole Hanks, Robert Cole, Elizabeth Anson, Kimberly Sidora-Arcoleo, Dennis W. Luckey, et al. 2007. Effects of Nurse Home Visiting on Maternal and Child Functioning: Age-9 Follow-up of a Randomized Trial. *Pediatrics* 120:832–45.
- Runkel, Daniela. 2009. *EFFEKT-Interkulturell: Implementierung und Evaluation eines präventiven Eltern- und Kindertrainings an Grundschulen mit einem hohen Anteil von Kindern mit Migrationshintergrund*. PhD diss., Universität Erlangen-Nürnberg.
- Runkel, Daniela, Friedrich Lösel, Mark Stemmler, and Stefanie Jaursch. Forthcoming. Preventing Social Behaviour Problems in Children From Deprived Migrant Families: Evaluation of a Child and Parent Training.
- Shelton, Karen K., Paul J. Frick, and Jane Wootton. 1996. Assessment of Parenting Practices in Families of Elementary School-age Children. *Journal of Clinical Child Psychology* 25:317–29.
- Shure, Myrna B. 1990. *The PIPS Test Manual: A Cognitive Measure of Interpersonal Problem-Solving Ability for Four to Six Year Old Children*, 2d ed. Philadelphia: Hahneman University, Department of Mental Health Sciences.
- Shure, Myrna B. 1992. *I Can Problem Solve: An Interpersonal Cognitive Problem-Solving Program: Kindergarten & Primary Grades*. Champaign, IL: Research Press.
- Spivack, George, and Myrna B. Shure. 1982. Interpersonal cognitive problem solving and clinical theory. In *Advances in Child Clinical Psychology*, vol. 5, ed. B. Lahey and Alan E. Kazdin, 323–72). New York: Plenum.
- Stemmler, Mark, Andreas Beelmann, Stefanie Jaursch, and Friedrich Lösel. 2007. Improving Parenting Practices in Order to Prevent Child Behavior Problems: A Study on Parent Training as Part of the EFFEKT Program. *International Journal of Hygiene and Environmental Health* 210:563–70.
- Stemmler, Mark, Charlotte Kötter, Anneke Bühler, Stefanie Jaursch, Andreas Beelmann, and Friedrich Lösel. Forthcoming. Prevention of Familial Transmission of Depression Through a Family Oriented Programme Targeting Parenting as Well as the Child's Social Competence. *Journal of Children's Services*.
- Stemmler, Mark, Friedrich Lösel, Andreas Beelmann, Stefanie Jaursch, and Barbara Zenkert. 2005. Child Problem Behavior in Kindergarten and in Primary School: A Comparison Between Prediction Configural Frequency Analysis and Multiple Regression. *Psychology Science* 47 (3/4): 467–78.
- Steyer, Rolf, Siegfried Gabler, Alina A. von Davier, Christof Nachtigall, and Thomas Buhl. 2000. Causal Regression Models I: Individual and Average Causal Effects. *Methods of Psychological Research Online* 5:39–71.
- Tremblay, Richard E., Linda Pagani-Kurtz, Frank Vitaro, Louise C. Masse, and Robert O. Pihl. 1995. A Bimodal Preventive Intervention for Disruptive Kindergarten Boys: Its Impact Through Mid-Adolescence. *Journal of Consulting and Clinical Psychology* 63:560–68.
- Tremblay, Richard E., Frank Vitaro, Claude Gagnon, C. Piché, and N. Royer. 1992. A Prosocial Scale for the Preschool Behavior Questionnaire: Concurrent and Predictive Correlates. *International Journal of Behavioral Development* 15:227–45.
- Vitaro, Frank, Mara Brendgen, and Richard E. Tremblay. 2001. Preventive Intervention: Assessing its Effects on the Trajectories of Delinquency and Testing for Mediation Processes. *Applied Developmental Science* 5:201–13.
- Webster-Stratton, Carolyn, M., John Reid, and Mary Hammond. 2004. Treating Children with Early-Onset Conduct Problems: Intervention Outcomes for Parent, Child, and Teacher Training. *Journal of Clinical Child and Adolescent Psychology* 33:105–24.