

Child social skills training in developmental crime prevention: Effects on antisocial behavior and social competence

Andreas Beelmann and Friedrich Lösel*
University of Jena (Germany) and * University of Cambridge (UK)

Social skills training for children is becoming increasingly popular as a measure for developmental crime prevention. Although previous reviews of such programs have shown positive effects, they have also revealed problems of research design, outcome measures, and long-term follow up. Accordingly, this article reports on a recent meta-analysis of randomized evaluations of the effect of social skills training in preventing antisocial behavior and promoting social competence. Of 841 retrievable references, 84 research reports with a total of 136 treatment-control comparisons fulfilled the eligibility criteria. Results showed a small but significant overall positive effect of $d = .39$ at post-intervention and $d = .28$ at follow-up (3 months and later). Effect sizes were somewhat greater for outcome measures of social competence than for measures of antisocial behavior, particularly when delinquency was assessed. Cognitive-behavioral programs revealed the best results in terms of generalization over time and on outcome criteria. In addition, prevention measures indicated for children and adolescents who already manifested some behavioral problems had higher effect sizes than universal approaches. Because most studies dealt with small sample sizes, non-official outcome data, and measurements after less than one year, the results should be interpreted with caution. Further high-quality studies with long-term empirical outcome criteria are needed, particularly outside the United States.

El entrenamiento en habilidades sociales en la prevención temprana de la delincuencia: los efectos en la conducta antisocial y la competencia social. El entrenamiento en habilidades sociales para niños es una estrategia cada vez más popular como una medida de prevención de la delincuencia. Este artículo presenta los resultados de un metaanálisis que toma en consideración diseños experimentales aleatorizados donde se empleó la técnica de habilidades sociales en el sentido indicado. Los resultados de 136 comparaciones entre grupos de tratamiento y grupos control mostraron un efecto positivo significativo aunque pequeño de $d = .39$ en la posintervención y de $d = .28$ en el seguimiento (tres meses o más). Los tamaños del efecto fueron algo mayores para las medidas de resultado de competencia social que para medidas de conducta antisocial, en particular cuando se evaluó de forma específica la conducta delictiva. Los programas cognitivoconductuales fueron los más destacados. Además, los resultados fueron mejores si los programas se orientaban a niños y jóvenes que ya presentaban problemas de conducta. No obstante, estos resultados han de interpretarse con cautela debido a diferentes problemas metodológicos, como muestras pequeñas, datos de delincuencia no oficiales y períodos de seguimiento de menos de un año.

A lack of social competencies is a frequently cited characteristic of aggressive and delinquent children and adolescents (e.g., Coie and Dodge, 1998; Farrington and Loeber, 2001; Frick, 1998; Lösel and Bender, 2003). Although the term social competence is often applied in different ways (see Caldarella and Merrell, 1997), there is clear empirical evidence that problematic modes of social information processing (Crick and Dodge, 1994; Gifford-Smith and Rabiner, 2004; Lösel and Bliesener, 2003), deficiency in social

problem solving abilities (Matthys and Lochman, 2005), and problematic peer relations (Bagwell, 2004; Parker and Asher 1987; Thornberry, 1998) play important roles in the development and continuation of antisocial careers.

Along with other prevention approaches such as parent training or school and community programs (see Farrington and Coid, 2003; Sherman, Gottfredson, MacKenzie et al, 1997; Tremblay and Craig, 1995; Wasserman and Miller, 1998), these results have formed the basis for the development and implementation of numerous programs for preventing antisocial behavior and promoting social behavior in children. Frequently called social skills training, these approaches aim to promote behavioral competencies such as asking for assistance or offering invitations and social-cognitive skills such as nonaggressive modes of perception and attribution in ambivalent social situations, the ability to deal with problems in interpersonal interactions, or

effective strategies to control aggressive and violent behavioral impulses (e.g., Greenberg et al, 1995; Lochman et al, 1993; Spence, 2003; Webster-Stratton and Hammond, 1997). Typically, such programs contain a sequence of manual-based lessons that are delivered in a group format at pre-school or school (e.g., Lösel and Beelmann, 2005).

The effectiveness of social skills training has been the topic of several reviews and meta-analyses (e.g., Ang and Hughes, 2002; Beelmann, Pfingsten and Lösel, 1994; Kazdin, 1997; Schneider, 1992; Wilson, Gottfredson and Najaka, 2001; Wilson, Lipsey and Derzon, 2003). In sum, such reviews suggest social skills training is an effective form of intervention for preventing and treating behavioral problems in childhood and adolescence. Nonetheless, these and other authors emphasize the basic problems and particular deficiencies in research on the prevention of antisocial behavior (e.g., Bullis, Walker and Sprague, 2001; Gresham, 1998; Taylor, Eddy and Biglan, 1999). Our own earlier analysis showed that whereas effects are medium to large in proximal criteria (e.g., social-problem solving), they rarely generalize to broader constructs (e.g., peer acceptance) and everyday behavior as assessed by, for example, teacher reports (Beelmann et al, 1994). Other frequent problems are small sample sizes, weak designs, and a lack of long-term follow-up periods. All this raises questions about whether social skills training for young people is already a proven measure in preventing antisocial development and later offending. One must also emphasize that the oft-quoted studies on long-term effects of developmental crime prevention addressed more complex and intensive multi-modal family-oriented programs (e.g., Olds et al, 1998; Schweinhart, Barnes and Weikart, 1993; Tremblay et al, 1995; for an overview: Farrington and Welsh, 2003). These should not be confused with mostly relatively short and child-focused social skills training programs.

Against this background, the Crime and Justice Group of the Campbell Collaboration (Farrington and Petrosino, 2001, Garrido, Farrington and Welsh, this issue) launched a systematic, comprehensive, and up-to-date meta-analysis of methodologically sound studies on the beneficial effects of child social skills training in preventing antisocial behavior and criminality. This article is an intermediate report on the work of this review and summarizes the methods used and some of the results. Further results and a more comprehensive report on methodological aspects of this work are reported in Lösel and Beelmann (2003a, 2005, 2006).¹

Method

Study Selection

Primary studies were selected according to the following six criteria: (1) A study had to contain an evaluation addressing only a *social skills training program for the prevention of antisocial behavior in children and adolescents*. We excluded all studies evaluating programs with additional components (e.g., parent training, teacher training, or home visits). Likewise, we did not include programs focusing on other areas such as internalizing problems, drug prevention, and coping with divorce. (2) A study had to compare a *treatment and a control group in a randomized experimental design*. Although quasi-experiments were excluded in principle, we did include stratified modes of randomization (e.g., randomized field trial, randomized block design, matching plus randomization). Pre- and post-intervention data also had to be

available. (3) the *age* of the youngsters treated had to be between 0 and 18 years. (4) Although the focus was on primary *prevention* programs (universal or targeted) rather than on clinical treatment, we included secondary prevention programs for youngsters with conduct disorders or oppositional-defiant disorders, because these are specific risk groups for later offending. However, treatment programs for delinquents already adjudicated were excluded. (5) The studies had to report *outcomes of a measure of antisocial behavior* (e.g., aggression, delinquency, disruption, or other antisocial behavior) and/or data on *social competence* (e.g., social interaction skills, pro-social behavior, or specific social-cognitive skills such as self-control or social problem-solving skills). A broad range of data sources was included (e.g., self-reports, parent reports, teacher reports, peer reports, observational data, and official records). However, data had to be reported in sufficient detail to permit reliable effect size computation. Finally, (6) we included all retrievable published or unpublished reports in *English or German* that appeared not later than the year 2000.

Literature Search

Several search strategies were applied to identify relevant evaluation studies: First, electronic databases such as *Psychinfo*, *Medline*, *Eric*, and *Dissertation Abstracts* were searched intensively. Second, the references from reviews on child skills training and the prevention of antisocial behavior were checked systematically. Third, the references given in previously identified primary studies were analyzed for further relevant publications.

A total of 851 articles were identified in this way. From these, 230 reports were excluded in a first screening because they obviously did not fulfill the selection criteria. The remaining 621 articles (80% published and 20% unpublished) were checked in more detail (see Lösel and Beelmann, 2003a). By excluding studies step by step, we ended up with 84 research reports (see Appendix) that met our eligibility criteria. Because a number of reports contained more than one treatment or control group or separate analyses for children and adolescents or for boys and girls, the final database (and our unit of analysis) for this meta-analysis was 136 treatment-control-group comparisons. Altogether, these contained 16,723 youngsters (of whom 52.1% belonged to the treatment groups).

Coding and Computation of Effect Sizes

The first author and a trained student coded all comparisons according to a detailed scheme. This contained characteristics of publication (e.g., year, country), methods (e.g., design, follow up), intervention programs (e.g., type, intensity, setting), and the children trained (e.g., age, gender, risk factors). A selection of these variables is presented in the results section (table 1). Two coders analyzed a subsample of 24 comparisons independently. Depending on category, interrater agreement varied between 81% and 100% ($M = 96.3\%$).

We used Cohen's (1988) *d* coefficient to compute unified effect sizes. When relevant data were available, we computed the effect sizes as the difference between the pre- to post-test (or follow-up) scores in the program group and the control group divided by the pooled standard deviation in the pretest. If no means and standard deviations had been reported, re-computation and effect size estimation techniques were used (see Lipsey and Wilson, 2001). If

the reports mentioned non-significant results with no details, we counted these as zero effects. In addition, we addressed the problem of outlier effect sizes (see Lipsey and Wilson, 2001). On the one hand, outliers are assumed to have questionable reliability with, at times, a strong influence on mean effect size. On the other hand, the high positive effect in these cases should not be ignored totally. Therefore, as a compromise, we restricted every effect size that exceeded the value of three standard deviations to 3.0 ($n= 6$).

Integration and Statistical Analysis

In a number of studies, the post-intervention measures were not assessed immediately after the training but several months later. Other studies had follow-up periods that were shorter than such post-tests. To solve this problem, we used a common time metric to construct unified measurement periods. Due to the small number of studies with relatively long follow-up assessments (see table 1), we used only two categories: Each effect size referring to assessments up to 3 months after training was categorized as a «post-intervention» effect. Measures that had been assessed later were subsumed to the «follow-up» category. This strategy produced 509 individual post-intervention and 117 individual follow-up effect sizes. From these effect sizes, we computed the study (contrast) effect size in a two-step procedure. First, each individual effect size within studies was integrated by calculating the mean effect within the two categories of outcomes (antisocial behavior, social competence). Second, we calculated the mean study effect by integrating effect size across the two broad categories. Accordingly, there was only one effect size for each category and for each treatment-control-contrast at the two measurement times. This method allowed us to perform independent analyses of overall study effects and of both outcome categories.

Mean effect sizes across treatment-control contrasts were calculated according to methods proposed by Hedges and Olkin (1985). These included weighting effect sizes by the inverse of sampling error and performing subsequent homogeneity analyses in order to analyze effect size variance (e.g., whether effect size variance exceeds sampling error). When effect sizes were homogeneous, the fixed effect model was applied, when they were heterogeneous, the random effect model was used (see, for details, Lipsey and Wilson, 2001).

Results

Description of Studies and Comparisons

Studies were published between 1971 and 2000 and—with one exception (Beelmann, 2000)—were written in English. Most studies during the last two decades were conducted in the United States ($k= 71$, 84.5%) and were published in scientific journals ($k= 78$, 92.9%). The low rate of unpublished reports ($k= 4$, 4.8%) may have been due to our restriction to randomized trials. A further description of the 136 treatment-control contrasts is given in table 1.

Although we selected studies with a relatively high quality of study design, other methodological aspects still reflected some of the major problems facing evaluation research in general. As in other treatment areas, most contrasts were based on only small samples of less than 50 participants (72.6%). In addition, about 56.3% had no follow-up assessment. Furthermore, when measuring the stability of effects, follow-up intervals were relatively short.

Study Characteristics	Coding	Frequency	Percent
General study characteristics			
Publication year ^a	Up to 1980	19	22.6
	1981-1990	39	46.4
	1991-2000	26	31.0
Publication type ^a	Journal article	78	92.9
	Book, Chapter	2	2.4
	Unpublished	4	4.8
Country ^a	USA	71	84.5
	Canada	8	9.5
	Other	5	5.9
Methodological characteristics			
Sample size	< 30	58	42.6
	30-49	43	31.6
	50-149	16	11.8
	150-500	15	11.0
	> 500	4	2.9
Type of outcome comparison ^b	Postintervention only	102	75.0
	Post and follow-up	25	18.4
	Follow-up only	9	6.6
Time at latest outcome measurement	Up to 1 month	50	60.7
	1 - 2 months	5	4.5
	3 - 6 months ^c	17	12.3
	12 months ^c	11	12.3
	> 12 months ^c	6	5.6
Treatment characteristics			
Type of treatment	Behavioral	38	27.7
	Cognitive	29	21.3
	Cognitive-behavioral	48	35.3
	Counseling, psychotherapy, etc	21	15.4
Number of sessions	Up to 10	56	44.8
	11-30	46	33.8
	31-60	22	16.2
	> 100	1	0.7
	Not specified	11	8.1
Treatment duration	Up to 1 month	24	18.0
	1-2 months	46	31.8
	2-4 months	38	27.9
	4-6 months	12	8.8
	6-12 months	11	8.1
	> 12 months	2	1.5
Trainer	Teachers	32	23.5
	Psycho-social professionals	35	25.7
	Study authors, Research staff	22	16.2
	Supervised students	30	22.1
	Others	4	2.9
	Not specified	13	9.6
Child characteristics			
Age (years)	4 - 6	26	19.1
	7 - 9	54	39.7
	10 - 12	38	27.9
	13 - 18	18	13.2
Gender (% male)	0	7	5.1
	40 - 59	44	32.4
	60 - 79	28	20.6
	80 - 99	18	13.2
	100	24	17.6
	Not specified	15	11.0
Type of prevention	Universal	31	22.8
	Selective	54	39.7
	Indicated	51	37.5

Note: ^a Based on 84 research reports. ^b Postintervention= All effects measured within two months after treatment. Follow up= All effects measured three months or more after treatment.

They ranged between 3 and 42 months with a mean of 6.9 months. Only five studies had measurements more than 12 months after the termination of training. The integrated studies used a relatively broad range of different dependent variables. At post-intervention, nearly 60% of the contrasts ($k= 82$) contained measurements of antisocial behavior, and these were based on 190 individual effect sizes (e.g., $M= 2.38$ effect sizes per contrast). At follow-up, twenty contrasts with 46 individual effect sizes measured in this area. Social competencies were assessed in 92 contrasts with 319 individual effect sizes (follow-up: 24 contrasts with 71 individual effect sizes, respectively). Taken together, each contrast had a mean of 4.10 (Range 1-19) individual effect measures at post-intervention and 3.68 (Range 1-10) at follow up.

Most programs had a behavioral and/or cognitive orientation. Combined approaches addressing both social-cognitive skills and concrete social behavior were most frequent. Other programs such as counseling, psychotherapy, or intensive care were investigated less frequently. Programs were generally relatively short interventions of limited intensity. Over 40% contained no more than 10 sessions and about half lasted no longer than 2 months. The typical training format was group training ($k= 105, 77.8%$) carried out in a school setting ($k= 100, 74.1%$). Nearly half of the programs were conducted by teachers or psychosocial professionals in practical contexts. In a comparable number of contrasts, the trainers were study authors, research staff, or supervised students.

The mean age of the children trained varied from 4 to 18 years. However, there was a clear focus on pre-school/elementary school age. More than 80% of the contrasts addressed children younger than 12 years. Although most studies contained mixed samples of boys and girls, boys were over-represented in general. This is in line with the higher prevalence of antisocial behavior in males. Programs targeting children who had already exhibited some form of antisocial behavior (indicated prevention) or who had other risk factors such as deficiency in social skills or academic achievement (selective prevention) were more frequent than programs for unselected «normal» groups (universal prevention).

Overall Intervention Effects

Post-intervention individual effect sizes ($n= 509$) ranged between -1.89 and 4.91 ($M= 0.39$). Although 16.1% of the effects were negative, the majority revealed a positive outcome (better results for the treated group). The unweighted overall mean for study effect size at post-intervention was $d= 0.48$ ($k= 127$). Follow-up individual effect sizes ranged between -2.39 and 4.33 ($M= 0.38$). The unweighted overall mean study effect size was $d= 0.38$ ($k= 34$). Table 2 shows the more adequate weighted mean effect size.

Following the fixed effect model, the overall effect became smaller when the treatment-control contrasts were integrated by weighting for sample size. This indicates smaller effects in larger samples. However, the total effect remained significant ($p<.05$) in both the fixed and the random models. Because the fixed model revealed significant heterogeneity beyond sampling error ($Q [df=126]= 283.23, p<.001$), the random model seemed to be more appropriate for our data (see Lipsey and Wilson, 2001). According to this model, the mean total post-intervention effect was $d= .39$ (equivalent to $r= .19$).

The mean effect size of the follow-up measurements was slightly smaller than that of the post-intervention measurements. However, effects still remained significant (i.e., $d= .28$, equivalent to $r= .14$) in the random model. Similar relations were found when we compared only those studies in which both post-intervention and follow-up measurements had been collected. Here, the mean overall effect was $d= .36$ in the post-test and $.32$ in the follow up (random model).

Moderator Analysis

Because both the overall effect size and the effect sizes for the two outcome criteria showed significant heterogeneity, we selected the random model of effect size integration for our moderator analysis (see table 3).

Despite some minor non-significant differences between the *types of treatment* in the overall effect size, cognitive-behavioral programs were the only category with significant effects on both antisocial behavior and social competence at post-intervention. For antisocial behavior, these differences between treatment types almost attained significance ($Q [df= 3]= 7.11, p<.06$). Follow-up analyses revealed the same pattern of results. Most importantly, cognitive-behavioral interventions were the only type with a significant influence on antisocial behavior ($d= .50$).

Treatment intensity was not a significant moderator at post-intervention or follow-up. However, there was a slight tendency for very intensive treatments to produce the highest effect size in both antisocial behavior and social competence. For example, at follow-up, highly intensive treatments were the only category with a significant effect on antisocial behavior ($d= .30$).

A statistical trend was found at post-intervention depending on the *category of trainer*. When authors, project staff, or supervised students served as trainers, effects were greater than when the program was conducted by teachers or other psychosocial practitioners ($d= .47$ vs. $.33$; $Q [df= 1]= 2.82, p<.09$). This result did not hold true for antisocial behavior, but was mainly attributable to measures of social competence ($d= .53$ vs. $.34$) that revealed a significant difference at post-intervention ($Q [df= 1]= 3.85, p<.05$).

Age of children was no significant moderator in the overall post-intervention outcome. However, this was not the case for measures of antisocial behavior, in which the oldest age group revealed the largest effect size ($d= .61, Q [df= 2]= 6.08, p<.05$). Moreover, programs for the youngest and oldest children showed the largest effect in the follow-up ($Q [df= 1]= 16.59, p<.001$). However, some of these findings were based on only a few studies (e.g., $k= 2$ for children 13 and older). When the various outcome criteria were considered, the only highly significant long-term effects in studies with 4- to 6-year-olds were for social competence ($d= .72$). In contrast, antisocial behavior was significant only for the oldest group ($d= .82$).

Table 2
Mean effects of child social skills trainings by measurement time and type of outcome

Outcome measure	Postintervention			Follow-up		
	Fixed model	Random model	k	Fixed model	Random model	k
Antisocial behavior	.19*	.29*	82	.06	.20*	20
Social competence	.32*	.43*	92	.16	.31*	24
TOTAL	.27*	.39*	127	.14*	.28*	34

Table 3
Relation between study characteristics and postintervention and follow-up effect sizes on antisocial behavior and social competence

Moderator	Postintervention						Follow up					
	Antisocial behavior		Social competence		Total		Antisocial behavior		Social competence		Total	
	d	k	d	k	d	k	d	k	d	k	d	k
Type of treatment												
Behavioral	.14	25	.50 ^{ab}	22	.34 ^a	37	.12	4	.34 ^a	4	.17	5
Cognitive	.14	15	.47 ^a	21	.41 ^a	25	-.06	3	.41 ^a	8	.36 ^{ab}	9
Cognitive-behavioral	.49 ^{ab}	26	.41 ^a	41	.43 ^a	48	.50 ^{ab}	7	.27 ^a	11	.37 ^{ab}	14
Other	.38 ^a	16	.30	8	.37 ^a	17	.16	6	.30 ^a	1	.17	6
Treatment intensity ^c												
low	.20	24	.43 ^a	39	.39 ^a	48	.12	3	.22 ^a	5	.22	7
medium	.31 ^{ab}	49	.42 ^a	49	.38 ^a	69	.17	11	.34 ^a	19	.31 ^a	21
high	.46 ^a	9	.60 ^{ab}	4	.46 ^{ab}	10	.30 ^a	6	–	–	.30 ^a	6
Trainers												
Teacher, professionals	.33 ^{ab}	44	.34 ^a	42	.33 ^a	60	.23 ^{ab}	13	.28 ^a	15	.27 ^a	21
Authors, staff, students	.24 ^{ab}	27	.53 ^a	36	.47 ^a	50	.20	5	.37 ^a	8	.34 ^a	10
Age												
4-6	.19	16	.43 ^a	18	.33 ^a	24	.12	1	.72 ^a	5	.60 ^a	6
7-12	.24 ^a	50	.40 ^a	61	.38 ^a	85	.17	17	.24 ^a	19	.22 ^a	26
13 and older	.61 ^{ab}	15	.67 ^a	13	.51 ^{ab}	18	.78 ^{ab}	2	–	–	.78 ^{ab}	2
Type of prevention												
Universal	.07	13	.45 ^a	23	.36 ^a	30	-.05	2	.14 ^a	4	.15	4
Selective	.12	27	.40 ^a	33	.31 ^a	46	.15	10	.30 ^a	10	.23 ^a	18
Indicated	.53 ^{ab}	42	.44 ^a	36	.49 ^{ab}	51	.48 ^{ab}	8	.40 ^a	10	.41 ^{ab}	12

a= Effect size differs significantly from zero ($p < .05$). b= Effect size shows significant heterogeneity ($p < .05$). c= Coding: Low= up to 10 sessions or 2 month duration, moderate= 11 to 40 sessions or 3 to 8 months duration, high= more than 40 sessions or 8 months duration.

Turning to the *type of prevention*, indicated programs tended to have the largest effects at both post-intervention and follow-up. Accordingly, the effect size was somewhat larger in groups with multiple risk factors than in groups from the general population. This pattern became most pronounced for antisocial behavior, in which only the indicated prevention programs had a significant effect size at both post-intervention and follow-up ($d = .53$ and $.48$). These effects were considerably higher than for universal and selective prevention (post-intervention: $Q [df = 2] = 13.25$; Follow-up: $Q [df = 2] = 4.84$, $p < .09$). The strongest effects of universal and selective strategies were found in measures of social competence, although universal preventive measures had no significant follow-up effect on either antisocial behavior or social competence.

Most further analyses of methodological and substantial moderators revealed non-significant results (see Lösel and Beelmann, 2003a, 2005, 2006). One important exception was the *sample size*, which had a significant and negative linear effect on outcomes ($Q [df = 2] = 6.20$, $p < .05$). Studies with small sample sizes (up to 30) had the highest ($d = .50$, $k = 66$, $p < .01$) and studies with sample sizes greater than 100 the lowest effect size ($d = .25$, $k = 15$, $p < .05$). Studies with sample sizes between 30 and 100 had intermediate effect sizes ($d = .34$, $k = 46$, $p < .05$). Findings were nearly identical for follow-up results and for effect sizes on antisocial behavior and social competence.

Discussion

Our review revealed a substantial number of randomized studies on the effect of social skills training on developmental prevention of antisocial behavior. Using the most appropriate computation model, 127 treatment-control contrasts with a total of over 16,500 youngsters yield a post-intervention effect of $d = .39$. This is equivalent to a correlation of $r = .19$ and thus a small effect size according to the terminology of Cohen (1988). It is somewhat lower than that in previous meta-analyses on social skills training in general (e.g., Beelmann et al, 1994; Schneider, 1992). This may be partially due to our specific focus on antisocial behavior and prevention programs as well as the restriction to particularly well-controlled (randomized) designs. In any case, there are several reasons why small effects should not be underestimated in practice. First, assuming that 50% of the control group would develop at least some temporary behavior problems, an effect size of $r = .19$ means that the treatment conditions show a reduction of 19 percentage points or 38% of these cases. Second, because many child social skills programs are relatively short and delivered in group settings, small to medium effects may well pay off in terms of cost effectiveness (see Welsh and Farrington, 2001). Third, social skills training programs often have practical advantages because they reach the whole target population and can be more easily and less expensively delivered than parent—or family—oriented programs (Offord et al, 1998; Prinz and Miller, 1994;

Taylor and Biglan, 1998). Fourth, a large body of research shows that the predictive power of most well-known risk factors for antisocial behavior and criminality is rarely larger than $r = .20$ (Hawkins et al, 1998; Lösel, 2002). When we consider that only a small number of risks can be addressed in child skills training programs and other risks may remain operative for years, our mean effect size seems to be plausible. Finally, the present meta-analysis permits relatively reliable conclusions because only studies with randomized control-group designs have been included.

Despite these positive aspects, several differential findings do not regard child skills training as an already well-proven measure for the prevention of antisocial developments: The follow-up effects are lower than the post-intervention outcomes and are rarely based on a time period of more than one year. In addition, effects on antisocial behavior are somewhat lower than those on social competence. Both findings indicate the need for caution regarding the long-term effects of social skills training on the prevention of criminal careers (see, McCord, 2003). In addition, most studies have only small sample sizes, and this also impacts on the outcome. As in other meta-analysis (e.g., Farrington & Welsh, 2003; Lipsey and Wilson, 1998), studies with small samples reveal the strongest post-intervention effects. To some extent, this may well reflect a selective publication of significant results (so-called publication bias).

However, larger effects in small studies may also be due to advantages in program implementation. For example, in large studies, difficulties in maintaining program integrity and the homogeneity of samples may reduce design sensitivity and thus lead to smaller effects (Lösel and Wittmann, 1989; Weisburd, Petrosino and Mason 1993). The relationship between quality of implementation and effectiveness may also be responsible for the larger effects (at least in social competence measures) when programs are delivered by the study authors, other research staff, or supervised students. Such trainers probably have a particularly strong interest in delivering the training as planned, whereas teachers and other practitioners may modify concepts to fit their real-life requirements and have to cope with less favorable circumstances than those in demonstration projects (see also Wilson, Lipsey and Derzon, 2003). These and other results indicate the need for more process evaluations and analyses of the implementation and integrity of programs in everyday practice (Greenberg, 2004).

The present results on child skills training are quite comparable with other measures of developmental prevention in this field. For example, Farrington and Welsh (2003) report a mean effect size of $d = .32$ on delinquency measures for parent- and family-oriented programs. Beelmann and Bogner (2005) have recently conducted a comprehensive meta-analysis on behavioral parent training programs. Although they report a relatively large general post-intervention effect of $d = .64$, effect sizes are much smaller for measures of antisocial behavior in follow-up assessments ($d = .29$).

The finding that both child- and parent-oriented programs have significant outcomes suggests that a combined multi-modal approach may help to increase effect size. This can be demonstrated by various studies (e.g., Kazdin et al, 1987; Lösel et al, 2005; Webster-Stratton and Hammond, 1997). Such cumulative effects (Lösel and Bender, 2003; Yoshikawa, 1994) are also confirmed by the positive long-term outcomes from some very comprehensive multi-modal early prevention programs (e.g., Olds et al, 1998; Schweinhardt et al, 1993, 2005; for a review, see

Tremblay and Japel, 2003). However, at times, even such comprehensive approaches produce only small effect sizes when subjected to rigorous evaluation (e.g., Conduct Problems Prevention Research Group, 2002). Thus, there is still a great need for further systematic program evaluations, particularly for studies that compare various programs or test effective program combinations (Greenberg, 2004).

Beside sample size, our meta-analysis reveals other, more content-related moderators of effectiveness. Although there is no significant effect of treatment type on the total outcome at post-intervention, differential results reveal a clearer picture. In particular, only cognitive-behavioral programs have significant effects on both post-intervention and follow-up measurements of antisocial behavior. This is in line with the finding that well-structured, multimodal, cognitive-behavioral programs are also relatively successful in the field of offender treatment (see Lösel, 1995; Lipsey, 2003; McGuire, 2001). However, in both fields the concrete program delivery seems to be just as important as the conceptual background and content of programs (Lösel, 2001).

Interestingly, treatment intensity is not a significant moderator in our meta-analysis. This may be because the typical program dosage is relatively low and does not vary that much across programs. Nonetheless, high-intensity treatments always have the highest effect sizes and are the only ones to reveal a significant long-term impact on antisocial behavior. This is in line with the need for more long-lasting programs to prevent antisocial careers in high-risk groups (e.g., Loeber and Farrington, 2001; Loeber and Stouthamer-Loeber, 1998; Lösel and Beelmann, 2003b).

Another finding from our research synthesis is that child characteristics correlate significantly with effect size variance. In particular, the largest effect sizes are found in studies on youth samples 13 years and older. At first glance, this result seems counterintuitive, and it conflicts with the idea that «earlier is better.» Common sense would lead us to expect the strongest program effects in young children (e.g., before behavioral problems can consolidate). In addition, programs addressing youngsters who have already developed some behavioral problems (indicated prevention) have the largest effects, whereas programs for general cohorts or unselected groups (universal prevention) have the lowest effects. However, a closer look on these findings makes them plausible: Although there may be positive short term effects in unselected groups, most of these youngsters would not develop serious behavioral problems in the long run even without the program. In addition, they already show at least «normal» levels of social competence and generally have a low level of behavioral problems. As a consequence, a large proportion of low-risk children cannot get much better through social skills training («ceiling effect»). This leads to non-significant mean differences in outcome behavior compared with untreated control groups. In contrast, in high-risk groups or groups that already exhibit behavioral problems, successful programs can have a stronger impact as is demonstrated in our evaluations of indicated prevention. Again, these findings are in line with research on delinquency treatment in which high-risk samples show larger effects (Lipsey and Wilson, 1998). More indirectly, they also seem to be in accordance with family-oriented delinquency prevention because clinic-based programs reveal larger effects than school-based programs (Farrington and Welsh, 2003).

Although universal programs are easier to implement and avoid problems of stigmatization (e.g., Offord et al, 1998;

Lösel, 2002), risk-focused programs may be more adequate for those who are most in need of prevention. However, this issue needs further clarification and a careful weighting of practical, legal, ethical, and financial issues (e.g., LeBlanc, 1998; Lösel, 2002). As a preliminary recommendation, child skills training may not so much be applied in large-scale implementations of universal prevention programs but more in risk-focused approaches for children and families with serious difficulties.

Our meta-analysis does not just provide a systematic review of program effects and moderators, but also reveals deficits in research. Alongside the lack of randomized studies with large samples and long follow-up periods, it is worth noting that only a few randomized studies have been performed outside the United States and Canada. This is a serious problem, because programs and findings cannot simply be generalized from one cultural context to another without further evaluation. We hope that the

international orientation of the Campbell Collaboration Group and this journal issue may help provoke an increase of high-quality evaluations outside North America.

Author's Note

The work on this study was supported by grants from the German Federal Ministry for the Family, Seniors, Woman, and Youth (BMFSFJ) and the German Research Council (DFG).

We thank Jonathan Harrow and Sara Harrop for English-language, native-speaker advice.

Footnote

¹ Due to differences in methodology and meta-analytic procedures (see below), the results reported here vary slightly from those in our previous publications.

References

- Ang, R.P. and Hughes, J.N. (2002). Differential benefits of skills training with antisocial youth based on group composition: a meta-analytic investigation. *School Psychology Review*, 33, 164-185.
- Bagwell, C.L. (2004). Friendships, peer networks and antisocial behavior. In J.B. Kupersmidt and K. A. Dodge (eds.): *Children's peer relations: from development to intervention* (pp. 37-57). Washington, DC: American Psychological Association.
- Beelmann, A. (2000). *Prävention dissozialer Entwicklungen: Psychologische Grundlagen und Evaluation früher kind- und familienbezogener Interventionsmaßnahmen [Prevention of antisocial development: Psychological foundations and evaluation of early child- and family-oriented interventions]*. Unpublished postdoctoral habilitation thesis, University of Erlangen-Nuremberg, Germany.
- Beelmann, A. and Bogner, J. (2005). *Die Effektivität von Elterntrainingsprogrammen. Eine Meta-Analyse zur Prävention und Behandlung dissozialer Verhaltensprobleme [The effectiveness of parent training programs: a meta-analysis on the prevention and treatment of antisocial behavior in children and adolescents]*. Paper presented at the Prevention Conference 2005 at Cologne, Germany, June 3rd – 4th, 2005.
- Beelmann, A., Pfungsten, U. and Lösel, F. (1994). The effects of training social competence in children: a meta-analysis of recent evaluation studies. *Journal of Clinical Child Psychology*, 23, 260-271.
- Bullis, M., Walker, H.M. and Sprague, J.R. (2001). A promise unfulfilled: social skills training with at-risk and antisocial children and youth. *Exceptionality*, 9, 67-90.
- Caldarella, P. and Merrell, K.W. (1997). Common dimensions of social skills of children and adolescents: a taxonomy of positive behaviors. *School Psychology Review*, 26, 264-179.
- Cohen, J. (1988). *Statistical power analysis for the behavioral science*. New York: Academic Press.
- Coie, J.D. and Dodge, K.A. (1998). Aggression and antisocial behavior. In W. Damon and N. Eisenberg (eds.): *Handbook of child psychology: vol. 3. Social, emotional and personality development* (5th ed., pp. 779-862). New York: Wiley.
- 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, 553-567.
- Crick, N.R. and Dodge, K.A. (1994). A review and reformulation of social information-processing mechanisms in children's social adjustment. *Psychological Bulletin*, 115, 74-101.
- Farrington, D.P. and Coid, J.W. (eds.). (2003). *Early prevention of adult antisocial behaviour*. Cambridge, UK: Cambridge University Press.
- Farrington, D.P. and Loeber, R. (2001). Summary of key conclusions. In R. Loeber and D.P. Farrington (eds.): *Child delinquents: development, intervention and service needs*. Thousand Oaks, CA.: Sage.
- Farrington, D.P. and Petrosino, A. (2001). The Campbell Collaboration Crime and Justice Group. *Annals of the American Academy of Political and Social Science*, 578, 35-49.
- Farrington, D.P. and Welsh, B.C. (2003). Family-based prevention of offending: a meta-analysis. *Australian and New Zealand Journal of Criminology*, 36, 127-151.
- Frick, P. J. (1998). *Conduct disorders and severe antisocial behavior*. New York: Plenum.
- Gifford-Smith, M.E. and Rabiner, D.L. (2004). Social information processing and children's social adjustment. In J.B. Kupersmidt and K.A. Dodge (eds.): *Children's peer relations: from development to intervention* (pp. 61-79). Washington, DC: American Psychological Association.
- Greenberg, M.T. (2004). Current and future challenges in school-based prevention: the researcher perspective. *Prevention Science*, 5, 5-13.
- Greenberg, M.T., Kusché, C.A., Cook, E.T. and Quamma, J.P. (1995). Promoting emotional competence in school-aged children: the affects of the PATHS curriculum. *Development and Psychopathology*, 7, 117-136.
- Gresham, F.M. (1998). Social skills training: Should we raze, remodel or rebuild? *Behavioral Disorders*, 24, 19-25.
- Hawkins, J.D., Herrenkohl, T., Farrington, D.P., Brewer, D., Catalano, R. F. and Harachi, T.W. (1998). A review of predictors of youth violence. In R. Loeber and D.P. Farrington (eds.): *Serious & violent juvenile offenders. Risk factors and successful interventions* (pp. 106-146). Thousand Oaks, CA: Sage.
- Hedges, L.V. and Olkin, I. (1985). *Statistical methods for meta-analysis*. New York: Academic Press.
- Kazdin, A.E. (1997). Practitioner review: psychosocial treatment for conduct disorder in children. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, 38, 161-178.
- Kazdin, A.E., Esveldt-Dawson, K., French, K. and Unis, A. (1987). Problem-solving skills training & relationship therapy in the treatment of antisocial child behavior. *Journal of Consulting and Clinical Psychology*, 55, 76-85.
- Le Blanc, M. (1998). Screening of serious and violent juvenile offenders. In R. Loeber and D.P. Farrington (eds.): *Serious and Violent Juvenile Offenders: Risk Factors and Successful Interventions*. Thousand Oaks, CA: Sage.
- Lipsey, M.W. and Landenberger, N.A. (2003). Cognitive-behavioral interventions. In B.C. Welsh and D.P. Farrington (eds.): *Preventing crime: What works for children, offenders, victims and places*. Dordrecht, The Netherlands: Springer.

- Lipsey, M.W. and Wilson, D.B. (1998). Effective intervention for serious juvenile offenders: a synthesis of research. In R. Loeber and D.P. Farrington (eds.): *Serious and violent juvenile offenders. Risk factors and successful interventions* (pp. 313-345). Thousand Oaks, CA: Sage.
- Lipsey, M.W. and Wilson, D.B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage.
- Lochmann, J.E., Coie, J.D., Underwood, M.K. and Terry, R. (1993). Effectiveness of a social relations intervention program for aggressive and nonaggressive, rejected children. *Journal of Consulting and Clinical Psychology, 61*, 1.053-1.058.
- Loeber, R. and Farrington, D.P. (eds.). (2001). *Child delinquents: development, interventions and service needs*. Thousand Oaks, CA: Sage.
- Loeber, R. and Stouthamer-Loeber, M. (1998). Development of juvenile aggression and violence. *American Psychologist, 53*, 242-259.
- Lösel, F. (1995). The efficacy of correctional treatment: a review and synthesis of meta-evaluations. In J. McGuire (ed.): *What works: reducing reoffending* (pp. 70-111). Chichester: Wiley.
- Lösel, F. (2001). Evaluating the effectiveness of correctional programs: bridging the gap between research and practice. In G.A. Bernfeld, D.P. Farrington and A.W. Leschied (eds.): *Offender rehabilitation in practice* (pp. 67-92). Chichester, UK: Wiley.
- Lösel, F. (2002). Risk/need assessment and prevention of antisocial development in young people: basic issues from a perspective of cautionary optimism. In R. Corrado, R. Roesch, S.D. Hart and J. Gierowski (eds.): *Multiproblem violent youth* (pp. 35-57). Amsterdam, NL: IOS Press/NATO Series.
- Lösel, F. and Beelmann, A. (2003a). 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, F. and Beelmann, A. (2003b). Early developmental prevention of aggression and delinquency. In F. Dünkel and K. Drenkhahn (eds.): *Youth violence: new patterns and local responses. Experiences in East and West* (pp. 220-244). Mönchengladbach, Germany: Forum.
- Lösel, F. and Beelmann, A. (2005). Social problem-solving programs for preventing antisocial behavior in children and youth. In M. McMurrin and J. McGuire (eds.): *Social problem solving and offending* (pp. 127-143). Chichester, UK: Wiley.
- Lösel, F. and Beelmann, A. (2006). Child skills training. In B.C. Welsh and D.P. Farrington (eds.): *Preventing crime: What works for children, offenders, victims and places* (pp. 33-54). Dordrecht, NL: Wadsworth Publishing.
- Lösel, F., Beelmann, A., Jaursch, S., Koglin, U. and Stemmler, M. (2005). Entwicklung und Prävention früher Probleme des Sozialverhaltens: Die Erlangen-Nürnberger Studie [Development and prevention of early social behavior problems: The Erlangen-Nuremberg Study]. In M. Cierpka (Hrsg.): *Möglichkeiten der Gewaltprävention* (S. 201-229). Göttingen, Germany: Vandenhoeck & Ruprecht.
- Lösel, F. and Bender, D. (2003). Resilience and protective factors. In D.P. Farrington and J. Coid (eds.): *Prevention of adult antisocial behavior* (pp. 130-204). Cambridge, UK: Cambridge University Press.
- Lösel, F. and Bliesener, T. (2003). *Aggression und Delinquenz unter Jugendlichen: Eine Studie über kognitive und soziale Bedingungen* [Aggression and delinquency in adolescence: a study on cognitive and social origins]. Neuwied: Luchterhand.
- Lösel, F. and Wittmann, W.W. (1989). The relationship of treatment integrity and intensity to outcome criteria. *New Directions for Program Evaluation, 42*, 97-108.
- Matthys, W. and Lochman, J.E. (2005). Social problem solving in aggressive children. In M. McMurrin and J. McGuire (eds.): *Social problem solving and offending* (pp. 51-66). Chichester: Wiley.
- McCord, J. (2003). Cures that harm: unanticipated outcomes of crime prevention programs. *Annals of the American Academy of Political and Social Science, 587*, 16-30.
- McGuire, J. (2001). What works in correctional intervention? Evidence and practical implications. In G.A. Bernfeld, D.P. Farrington and A.W. Leschied (eds.): *Offender rehabilitation in practice* (pp. 25-43). Chichester, UK: Wiley.
- Offord, D.R., Kraemer, H.C., Kazdin, A.E., Jensen, P. and Harrington, R. (1998). Lowering the burden of suffering from child psychiatric disorder. Trade-offs among clinical, targeted and universal interventions. *Journal of the American Academy of Child and Adolescent Psychiatry, 37*, 686-694.
- Olds, D.L., Henderson, C.R., Cole, R., Eckenrode, J., Kitzman, H., Luckey, D., Pettitt, L., Sidor, K., Morris, P. and Powers, J. (1998). Long-term effects of nurse home visitation on children's criminal and antisocial behavior. *Journal of the American Medical Association, 280*, 1.238-1.244.
- Parker, J.G. and Asher, S.R. (1987). Peer relations and later personal adjustment: are low accepted children at risk? *Psychological Bulletin, 102*(3), 357-389.
- Prinz, R.J. and Miller, G.E. (1994). Family-based treatment for childhood antisocial behavior: experimental influences on dropout and engagement. *Journal of Consulting and Clinical Psychology, 62*, 645-650.
- Schneider, B.H. (1992). Didactic methods for enhancing children's peer relations: a quantitative review. *Clinical Psychology Review, 12*, 363-382.
- Schweinhart, L.J., Montie, J., Xiang, Z., Barnett, W.S., Belfield, C.R. and Nores, M. (2005). *Lifetime effects: the High/Scope Perry Preschool study through age 40*. Ypsilanti, MI: High/Scope Press.
- Schweinhart, L.J., Barnes, H.V. and Weikart, D.P. (1993). *Significant benefits. The High/Scope Perry Preschool study through age 27*. Ypsilanti, MI: High/Scope Press.
- Sherman, L., Gottfredson, D., MacKenzie, D., Eck, J., Reuter, P. and Bushway, S. (1997). *Preventing crime: what works, what doesn't, what's promising*. Report to the United States Congress: University of Maryland.
- Spence, S.H. (2003). Social skills training with children and young people: theory, evidence and practice. *Child and Adolescent Mental Health, 8*, 84-96.
- Taylor, T.K., Eddy, J.M. and Biglan, A. (1999). Interpersonal skill training to reduce aggressive and delinquent behavior: limited evidence and the need for an evidence-based system of care. *Clinical Child and Family Psychology Review, 2*, 169-182.
- Taylor, T.K. and Biglan, A. (1998). Behavioral family interventions for improving child rearing: a review of the literature for clinicians and policy makers. *Clinical child and family psychology review, 1*, 41-60.
- Thornberry, T.P. (1998) Membership in youth gangs and involvement in serious and violent offending. In R. Loeber and D.P. Farrington (eds.): *Serious and Violent Juvenile Offenders: Risk Factors and Successful Interventions*. Thousand Oaks, CA: Sage.
- Tremblay, R.E. and Craig, W.M. (1995). Developmental crime prevention. In M. Tonry and D.P. Farrington (eds.): *Building a safer society: strategic approaches to crime prevention* (vol. 19, pp. 151-236). Chicago, IL: The University of Chicago Press.
- Tremblay, R.E., Pagani-Kurtz, L., Vitaro, F., Masse, L.C. and Pihl, R.O. (1995). A bimodal preventive intervention for disruptive kindergarten boys: its impact through mid-adolescence. *Journal of Consulting and Clinical Psychology, 63*, 560-568.
- Tremblay, R.E. and Japel, C. (2003). Prevention during pregnancy, infancy and the preschool years. In D.P. Farrington and J.W. Coid (eds.): *Early prevention of adult antisocial behaviour* (pp. 205-242). Cambridge: Cambridge University Press.
- Wasserman, G.A. and Miller, L.S. (1998). The prevention of serious and violent juvenile offending. In R. Loeber and D. P. Farrington (eds.): *Serious & violent juvenile offenders. Risk factors and successful interventions* (pp. 197-247). Thousand Oaks, CA: Sage.
- Webster-Stratton, C. and Hammond, M. (1997). Treating children with early-onset conduct problems: a comparison of child and parent training interventions. *Journal of Consulting and Clinical Psychology, 65*, 93-109.
- Weisburd, D., Petrosino, A. and Mason, G. (1993). Design sensitivity in criminal justice experiments. *Crime and Justice, 17*, 337-380.
- Welsh, B.C. and Farrington, D.P. (2001). A review of research on the monetary value of preventing crime. In B.C. Welsh, D.P. Farrington and L.W. Sherman (eds.): *Costs and benefits of preventing crime* (pp. 87-122). Oxford, UK: Westview Press.
- Wilson, D.B., Gottfredson, D.C. and Najaka, S.S. (2001). School-based prevention of problem behaviors: a meta-analysis. *Journal of Quantitative Criminology, 17*, 247-272.
- Wilson, S.J., Lipsey, M.W. and Derzon, J.H. (2003). The effects of school-based intervention programs on aggressive behavior: a meta-analysis. *Journal of Consulting and Clinical Psychology, 71*, 136-149.
- Yoshikawa, H. (1994). Prevention as cumulative protection: effects of early family support and education on chronic delinquency and its risks. *Psychological Bulletin, 115*, 28-54.