

Experiences in Disseminating Evidence-Based Prevention Programs in a Real-World Setting

Inga Frantz^{1,3} · Mark Stemmler² · Kurt Hahlweg³ · Julia Plück⁴ · Nina Heinrichs^{1,3}

Published online: 2 April 2015
© Society for Prevention Research 2015

Abstract The primary aim of family-based prevention programs is to promote children's health. Unfortunately, it is difficult to reach families with such evidence-based prevention programs (EBP). Therefore, implementing EBP on a population level could be a promising approach to reach more families, including those faced with socioeconomic challenges who are usually less likely to participate in randomized controlled trials (RCT). Is a population rollout appropriate to reach more and different families than those participating in RCT, especially those representative of the target population? We implemented three EBP in a city in an uncontrolled trial. The effects of this population rollout were tracked on the level of the participating families and on the level of all families living in the city. More than 3480 families (30 % of the population) with children up to 12 years of age participated based on practitioner report. Analyses indicate that a greater percentage of low socioeconomic-status families attended a program compared with a randomly surveyed sample from the city's general population. The sizes of the within-subject effect for parental strategies, child behavior problems, and children's

quality of life for a subsample of $n=411$ families were similar to those of other uncontrolled EBP studies. The study contributes to highly needed type 2 translation research. The population-based dissemination of EBP could be a promising approach to reach families at risk. However, there are considerable barriers to the implementation process, which currently limit the effectiveness of this rollout in a community.

Keywords Prevention · Dissemination · Implementation · Parenting · Child behavior

Approximately 15–20 % of all children develop emotional or behavioral problems (Belfer, 2008; National Research Council & Institute of Medicine, 2009). Although not all of these problems meet the threshold for clinical disorder, between 10 and 15 % of children are in need of treatment due to their emotional and/or behavioral problems. Child psychopathology comes along with other problems such as juvenile delinquency, substance abuse, and high costs for youth-welfare services. A number of well-researched prevention programs for families exist to prevent or minimize developmental difficulties and to support families and teachers in raising healthy children. These evidence-based prevention programs have often been examined under very structured research settings with high internal validity, but they have rarely been widely disseminated. Therefore, a very limited number of families are reached. In fact, most families in real-life settings are provided with untested primary prevention programs and early interventions that have unknown effects at least in Germany.

Electronic supplementary material The online version of this article (doi:10.1007/s11121-015-0554-y) contains supplementary material, which is available to authorized users.

✉ Inga Frantz
i.frantz@tu-braunschweig.de

¹ Department of Psychology, University of Bielefeld, Bielefeld, Germany

² Department of Psychology, University of Erlangen-Nuremberg, Erlangen-Nuremberg, Germany

³ Present address: Department of Psychology, University of Braunschweig, Braunschweig, Germany

⁴ Department of Child and Adolescent Psychiatry and Psychotherapy, Medical Faculty, University of Cologne, Cologne, Germany

Who Is (Not) Reached with Randomized Controlled Trials in Prevention and Early Intervention of Childhood Psychopathology?

Many research groups have reported that families with low socioeconomic status (SES) are among the hardest to reach with prevention programs (e.g., National Research Council &

Institute of Medicine, 2009). Authors argue that these families seem to suffer from multiple psychosocial stressors that make participation more difficult. They may also feel more stigmatized when approached directly or screened for “problems” (National Research Council & Institute of Medicine, 2009). Finally, low SES families also seem to make less use of external support (Welsh & Farrington, 2012). Accordingly, interventions that aim at reaching socially disadvantaged families need to be very easy to access. They should be non-stigmatizing, delivered in a “natural” setting (e.g., preschool) or for a “natural” group (e.g., every parent of a school class). These goals may be reached if prevention programs were offered at a population level. Previously, the programs offered should have been tested in a randomized controlled trial (RCT) to ensure significant effects. Thus, combining both population-level dissemination efforts and selecting evidence-based programs seems to be a promising approach. However, there are still a limited number of studies on this type 2 translation research (Spoth et al., 2013).

Population-Based Implementation of EBP

Lately, there have been efforts to implement family-based prevention programs on a large scale. Some studies implementing prevention programs on a large scale show positive results (e.g., Fagan, Hanson, Briney, & Hawkins, 2012; Olds, 2000; Prinz, Sanders, Shapiro, Whitaker, & Lutzker, 2009; Salmivalli & Poskiparta, 2012), while others report no or mixed effects (e.g., Eisner, Nagin, Ribeaud, & Malti, 2012). One potential reason for varying study outcomes might be different study foci (e.g., focus on efficacy, implementation, or dissemination). The present study aimed to track the effects of a naturalistic population-based implementation of prevention programs in a community setting. The following conceptual framework informed the implementation strategy employed:

1. A significant number of families are in need of help, but effective programs are rarely used in public health settings. Therefore, enhancing the use of evidence-based prevention programs in practice is in high demand (e.g., Spoth et al., 2013).
2. An implementation approach is needed that (1) enables efficient and widespread use of evidence-based prevention programs in practice, (2) reaches a large number of families, especially those at risk, and (3) has the capacity to be delivered widely throughout communities (e.g., Glasgow & Emmons, 2007).

We addressed those requirements as follows: (1) For a translation of evidence-based prevention in practice, we implemented the programs with an existing workforce of

practitioners working in family-service organizations. To ensure services were delivered efficiently, we chose programs that have proven their ability to be successfully applied in practice and implemented those with a staff-support system in place. (2) To increase the likelihood that families would participate, we offered programs across different prevention levels (universal and indicated) and consumers (i.e., parents, children, and teachers) as well as across organizations in the community (Glasgow & Emmons, 2007). (3) To facilitate the applicability of the implementation approach and the widespread use of the programs throughout communities, we integrated EBP in the existing infrastructure with limited resources, as communities’ financial capacity often is limited.

We assumed that offering these programs more widely, in the frame of a population approach in the existing family-service setting, would increase participation of families with low socioeconomic backgrounds. Furthermore, we assumed that programs for families would show similar effects on parents’ and children’s behaviors and child wellbeing compared with effects reported in other uncontrolled prevention trials. The present study therefore addresses three research questions: (1) How many families are reached in a population rollout and in what time?; (2) Do participating families (“consumer sample”) differ from families randomly drawn from the target population (“population sample”) in socioeconomic variables?; (3) What effect sizes will be reached in this uncontrolled natural study design? The current study extends the existing literature concerning real-world implementation of prevention programs in several unique ways. A naturalistic approach of high external validity is crucial, because translation barriers (e.g., limited funding, workload) make implementation of evidence-based prevention programs in routine health care settings more difficult (Glasgow & Emmons, 2007). Introducing programs in existing family-service infrastructure on site and engaging community partners in the study (Glasgow & Emmons, 2007; Fagan et al., 2012), while getting along with a limited budget, could be a promising approach to reduce contextual barriers to program use. By ongoing evaluation of the implementation process, we allow for further adaptation of the strategies applied. Finally, by comparing the outcomes of the consumer sample with those from other trials, we test whether the naturalistic approach may have the capacity to effectively reduce dysfunctional parenting and child behavior problems.

Method

Design

We selected one small-sized German city (Paderborn) based on city preference and feasibility of the implementation process on site. Phone interviews with randomly drawn

independent samples of families living in the city were conducted three times: pre-intervention (2010), 1 year (2011) and 2 years (2012) later (uncontrolled trial). The city has a population of about 147,000 inhabitants, with about 11,500 families with at least one child under the age of 13 (target population) living in the city. The study was part of a larger research project called FAMOS (e.g., Frantz & Heinrichs, 2015a, b).

Prevention Programs

We selected three prevention programs available in Germany that: (1) have at least one RCT published, with indication of positive results either on parenting behavior or on social competence in children and (2) were prepared to be disseminated on a larger scale (materials available, manualized, training procedures in place). We selected one child skills training (EFFEKT), one preschool teacher- and parent-based indicated prevention program (PEP), and one parenting program (Triple P).

EFFEKT

The EFFEKT training program is a manual-based group program in social problem-solving for preschool- and school-aged children (Lösel & Stemmler, 2012). It is based on the “I can problem solve” training by Shure (1992). The program was, however, modernized and modified for the German context. The course consists of two parts. The first part addresses basics such as elementary verbal concepts, identification of emotions, reflection on causality, and reasons for behavior. The second part contains the training of social cognitive problem-solving skills, such as providing alternative solutions in conflicts, anticipation of actions, and evaluation of consequences. Each of the 15 sessions lasts 30 to 60 min. The efficacy of the program has been evaluated (e.g., Lösel & Stemmler, 2012).

PEP

The Prevention Program for Externalizing Problem Behavior (PEP; Plueck et al., 2014) is based on published prevention and treatment manuals for children with disruptive behavior problems (e.g., McMahon & Forehand, 2003). It is an indicated prevention program for parents and preschool teachers of children from 3 to 6 years of age with hyperkinetic and oppositional problem behavior. The intervention comprises ten sessions, each lasting 90 to 120 min with five to six participants (parents or teachers) per group. It addresses parenting strategies to enhance a positive parent/teacher-child relationship as well as skills to manage disruptive child behavior. The training has been found to be effective (e.g., Plueck et al., 2014).

Triple P

The Triple P-Positive Parenting Program is a multi-level, preventively oriented parenting and family-support strategy (Sanders, Kirby, Tellegan, & Day, 2014). The program aims to prevent behavioral, emotional, and developmental problems in children by enhancing the knowledge, skills, and confidence of parents. It incorporates five levels of intervention for parents on a continuum of increasing strength. In the present study, Triple P levels 1 to 4 were used. *Level 1* is a universal parent information strategy, using media, tip sheets, and videotapes. *Level 2* provides early anticipatory developmental guidance for groups (seminar series) or individual parents (brief primary care). *Level 3*, a four-session intervention, includes active skills training for parents (primary care). *Level 4* is an intensive six-to-eight-session individual (standard) or group-directed parent training; it is also available for parents with special needs (e.g., stepping stones for parents of disabled children). RCT and meta-analyses demonstrate the effectiveness of the Triple P system (e.g., Sanders et al., 2014).

Procedures

In the installation phase, a full-time site coordinator was selected. He had comprehensive knowledge of the given family-service organizations in the community and coordinated the implementation. He also acted as an agent between evaluators, program staff, practitioners, families, and community stakeholders to enhance effective collaborations.

Assuming that 25 % of the population must be reached to make a difference at a population level (Glazemakers, 2012), training of 211 practitioners was funded, expecting that each facilitator would subsequently deliver at least two courses of the prevention program in his or her usual workplace (see Fig. A for more details, available online). In order to include a representative sample, the site coordinator recruited staff from multiple settings (public, church-related, private) and diverse institutions (e.g., schools, youth welfare). Staff-selection criteria were applied to ensure program use (see Fig. B, available online). However, there were no inclusion criteria regarding the staff’s work experience reflecting a typical sample. The practitioners received intensive, cost-free training during their regular working time. Afterwards, the staff delivered the services for families in his or her usual workplace. Staff provided courses of varying intensity and format (individual vs. group), in different settings (e.g., preschool, doctor’s office) for different groups (all children, parents, preschool teachers of a child with behavior problems) at different hours of the day. For practitioner support, the programs’ staff offered ongoing individual as well as group coaching for practitioners to ensure successful program delivery in the practice setting and to enable peer supervision. The

coordinator provided administrative staff support by, for example, fostering necessary changes in the organizations, providing technical assistance and providing program materials (see Fig. B for more details, available online). Treatment fidelity was assured by using manualized interventions tested in at least one RCT including high-quality training by program developers as well as ongoing coaching. We estimated the quality of the interventions provided by assessing the practitioners' self-efficacy in consultation skills (pre- and post-training, as well as 6, 12, 18, and 24 months later), the evaluation of the workshop, the amount of staff coaching, and consumer results. Furthermore, we made ongoing assessment of the staffs' program use and communicated the results from the 2012 assessment to the site coordinator, stakeholders, staff, and coaches to allow them to adapt and optimize staff support throughout the implementation process. To enhance knowledge of these services in families, the site coordinator organized a community-wide public-awareness campaign (see Fig. B for details, available online).

Participants

There are two "types" of participants: (1) the service providers ("facilitators") participating in the training for program delivery and (2) the families. The families may further be divided into those participating in one of the programs ("consumer sample") and those who were principally eligible to participate ("population sample").

Facilitators

Most of the practitioners were preschool teachers (54.1 %), social workers (24.2 %), or parent educators (8.7 %) and were employed in preschools (51.2 %), education and family services (22.5 %), youth-welfare services (10.0 %), or in private practice (8.6 %). They were mostly female (92.8 %), 40 years of age ($M=39.9$, $SD=10.4$) and had between 0 and 38 years ($M=15.7$, $SD=9.9$) of work experience in the field of family services.

Families

Level of Participating Families (Consumer Sample)

We assessed the locally provided number of courses for families via three sources of information: (1) practitioner self-report, (2) released program material for families, and (3) returned questionnaires from participating families. We sent self-report questionnaires to the facilitators 3, 6, 12, 18, and 24 months after training, asking for the number of families they had consulted since the training. For precise measurement of the total program use, all facilitators also reported

the number of provided courses for families at three reference-date assessments (2011, 2012, 2013). The site coordinator registered the amount of program material for families (e.g., workbooks for parents) he has released to the practitioners. The participating families ("consumer sample") reported on their parenting skills and child behavior problems before and after attending a course. Therefore, paper-pencil questionnaires were sent to the practitioners, who administered the forms to the participating families pre- and post-intervention. There were two variants of questionnaires, differing in the number and type of questionnaires included (one small package for low-intensity interventions, namely Triple P levels 2 (seminar series) and 3 and a more comprehensive package for all others). The decision to adapt the number of questionnaires collected to the intensity of the intervention delivered was based on the assumption that a full 15-page package of questionnaires would not have been appropriate and also would have been highly unlikely to be returned to the research group with low-intensity programs. Participants of Triple P levels 1 and 2 (brief primary care) received no questionnaires at all, because these interventions are impossible to identify as concrete, single events with a pre- and post-intervention assessment. For evaluating PEP and EFFEKT services, teacher report was also assessed. However, these will not be reported in the present study because of the focus on the population of parents and children in the current manuscript (see Table A for an overview of all assessed outcome measures, available online).

We received at least one questionnaire from each of 895 families. Complete parent-reported questionnaires were available for a total of $N=411$ families (45.9 %), representing the consumer sample of the present study. Two hundred ninety-eight were excluded because families participated in a low-level intervention (Triple P level 2 or 3) and therefore completed the small questionnaire package only; for $n=186$, only teacher report was available. Out of the 411 families, $n=174$ (42.4 %) participated in a Triple P intervention ($n=132$ group, $n=26$ standard, $n=16$ stepping stones), $n=38$ (9.2 %) in PEP, and $n=199$ (48.4 %) in EFFEKT ($n=125$ preschool, $n=74$ elementary school). In 85.3 % of the families, the mother completed the questionnaires; in 13.3 % the father; and in 1.4 % others (e.g., step-parents, foster parents). Mean age of the informant was 35.6 years ($SD=6.2$).

Random Sample of Eligible Families (Population Sample)

For the purpose of the present study, we decided to compare the consumer sample with the population sample assessed in 2011. The reasons for this decision were that (1) parenting behavior on the population level was only assessed in 2011 and 2012. This measure was added later in the study process based on Prof. Manfred Döpfner's (leader of the PEP workgroup) recommendation and applied for a randomly

selected third of the families. This was a compromise between not losing participants because of a very long phone interview and having the opportunity to compare the two samples (consumer and population samples) on parenting behavior; and (2) we chose the earlier assessment (2011 instead of 2012) for minimizing potential impact of the program implementation on population-level data. Thus, for the purpose of the present study, we used population data from one (2011) of the three assessments, while the consumer sample was assessed twice (pre- and post-intervention) between 2010 and 2013.

For the population sample recruitment, the city staff randomly selected 10 % of families of the local population with at least one child under the age of 13 (2011, $N=1,142$). This random sample could include families who had participated in one of the three programs (and therefore may have also been part of the consumer sample). In 2011, we contacted 82 % of the random sample (18 % non-relevant dropout, e.g., because no phone number was available); 54 % ($n=506$) participated and 46 % denied participation. Of the 506 families, $n=7$ were excluded because the child was older than expected and no longer covered by inclusion criteria. From the final population sample ($N=499$), parenting information was available for $n=149$. Most interviews were completed by mothers (85.0 %); in a minority of cases, the father (14.0 %) or others, for example, foster parents (1.0 %), answered the questions. Chi-square tests for independence as well as independent-sample t test indicated that the group of families answering the parenting questionnaire did not differ significantly from those who did not, regarding child behavior measures (Strengths and Difficulties Questionnaire (SDQ) $t(389)=-0.31$, $p=.76$; KINDL-R $t(388)=-0.77$, $p=.44$) and demographic statistics (age of child $t(497)=-1.07$, $p=.28$; gender of child $\chi^2(1, n=499)=0.11$, $p=.74$).

Measures

SES

The following indicators of SES were assessed in the population as well as in the consumer sample: parents' educational level (highest graduation), dependence on financial support from the state (receiving unemployment compensation, social welfare or housing benefits), and fathers' delinquency ("Did the child's father ever get in trouble with the law in his life?"). The migration background, defined by a child having at least one parent born in a country other than Germany, was assessed in the population sample only.

Parenting Behavior

To assess parenting skills, the short version of the German Parenting Scale (PS; Arnold et al. 1993; Naumann et al.,

2010) was administered. It measures dysfunctional parenting strategies with 13 items. Parents answer on a 7-point Likert scale to which degree they are likely to act in difficult parenting situations, with higher scores representing more dysfunctional parenting. To categorize the parenting skills, cutoffs for the total score were applied: ≥ 3.7 borderline and ≥ 4.0 clinical (Frantz & Heinrichs, 2014). The PS has adequate validity and reliability. In the current study, internal consistency was acceptable for the total score (pre, $\alpha=.77$; post, $\alpha=.68$) in the consumer sample, but rather poor in the population sample ($\alpha=.55$). This may be due to two different modes of administration (written questionnaire vs. administration over the phone).

The Positive Parenting Questionnaire (PPQ; adaption from Strayhorn & Weidman, 1988) assesses positive parental behavior (e.g., "I cuddle with my child") with 13 items. Parents report the frequency of their encouraging parenting on a 4-point Likert scale from 0 (*never*) to 3 (*very often*). Reliability for the total score was good in the consumer sample, $\alpha=.85$ (pre) and $\alpha=.85$ (post).

Child Behavior

Child behavior was assessed with two measures, the SDQ (Goodman, 1997) and the Social Behavior Questionnaire (SBQ; Tremblay, Vitaro, Gagnon, Piché, & Royer, 1992). The SDQ contains 25 items and consists of five subscales (emotional symptoms, conduct problems, hyperactivity, peer problems, prosocial behavior) as well as a total problem score (excluding the prosocial scale). The parents report on children's behavior on a 3-point Likert scale, with higher scores reflecting more behavior problems. Clinical cutoffs for the total score are available: borderline 14–16 and clinical ≥ 17 . In the current study, internal consistency for the total score was $\alpha=.81$ (pre) and $\alpha=.81$ (post) for the intervention and $\alpha=.82$ for the population sample.

Child behavior problems were also measured with the German adaption of the SBQ. Of the original 48-item questionnaire, only the subscales internalizing problems (8 items), conduct problems (12 items, consisting of aggressive behavior and destruction of property/delinquency), and impulsivity/attention problems (8 items) were administered. Parents reported on their child's behavior on a 3-point Likert scale, with higher scores representing more behavior problems. In the current study, reliability was good for the subscales conduct problems (pre, $\alpha=.87$; post, $\alpha=.85$), internalizing problems (pre, $\alpha=.83$; post, $\alpha=.82$), and attention problems (pre, $\alpha=.88$; post, $\alpha=.86$).

Children's health-related quality of life was assessed using the 24-item KINDL-R (Bullinger et al. 2008). It comprises five subscales (physical well-being, self-esteem, family, friends, and everyday functioning in (pre-)school) and a total score. Parents respond on a 5-point Likert scale. Scores can be

transformed to a value ranging from 0 to 100, with higher scores reflecting a better quality of life. In the present study, internal consistency for the total score was $\alpha=.82$ (pre) and $\alpha=.61$ (post) in the intervention sample and $\alpha=.81$ in the population sample.

Child behavior (SBQ, SDQ) and quality of life (KINDL-R) were assessed in children 3 years of age or older, as these questionnaires are not suitable for younger children.

Statistical Analyses

An alpha level of .05 was used for all statistical tests. Between-group effect size was calculated using pooled standard deviation.

In the consumer sample, missing data analysis (MVA) with SPSS was conducted, assessing the amount as well as the pattern of missing data. From the subsample of families with children from 0 to 2 years of age (only parent and demographic data, $n=17$), 20.8 % of the data was missing; for the subsample of children older than 2 years or with age as a missing value, 28.1 % of the data was missing. Little's MCAR test indicated that missingness was unrelated to the observed data for the subgroup of older children ($n=394$; $\chi^2=30,745.75$; $df=31$, 477 , $p=.998$) but not for the younger subgroup ($\chi^2=1037.99$; $df=332$, $p<.001$). Additionally, we examined if families with missing data on relevant sociodemographic variables (parents' educational level, delinquency of father, financial support from the state, informant) differ from those without missing data on outcome variable (SDQ, KINDL-R, SBQ, PS, PPQ) pre-intervention. *T* tests indicate that the groups did not differ significantly with two exceptions: Families with missing data on financial support ($n=54$, $M=2.3$, $SD=2.9$) or father's delinquency (missing: $n=56$, $M=2.4$, $SD=2.9$) reported less conduct problems compared with those with complete data (financial support: $n=340$, $M=3.4$, $SD=4.0$; $t(89.04)=-2.49$, $p=.02$, $d=-.28$; father's delinquency: $n=338$, $M=3.4$, $SD=4.0$; $t(94.58)=-2.43$, $p=.02$, $d=-.26$). Despite these exceptions, we decided to impute missing data in order to receive results of high external validity. We used EM algorithm for missing substitution, as recommended by Schafer and Graham (2002). In this procedure, new values for missing data are simulated on the basis of the observed data. To analyze the quality of the imputed datasets, we conducted *t* tests to investigate for potential differences between observed and imputed data. Outcomes indicated that imputed data did not differ from observed data (results available upon request from the first author). Therefore, only results of imputed data are presented.

In the population sample, a small percentage of the data was missing, 0.2 % for families with children from 0 to 2 years ($n=106$; demographic statistic and parent measure only) and 1.0 % for families with older children ($n=393$). Because of this very small percentage of missing data, we analyzed observed data only.

Results

Implementation

There was an increase in staff attrition over time: While 9 % of staff had dropped out until 1 year after training (2011), the figures were 17 % at 2012 assessment and 28 % at 2013 assessment. Nonetheless, there was an increase of practitioners that reported using the program over time: 53 % of practitioners had delivered at least one course for families at 2011 assessment, and this rate increased to 76 % (2012) and 82 % (2013).

The estimates of the total number of families reached revealed an increase of families reached over time: (1) Aggregated results of all assessments of practitioners' self-reported program use indicate that at least 1009 families (9 % of the referent population) were reached after 1 year (2011) and 2103 families (18 % of the population) after 2 years (2012). As a consequence of feed-backing, these results to the site coordinator and coaches, practitioner staff support was increased. At 2013 assessment, 3480 families (30 % of the population) had been reached. Thirty-seven percent ($n=1294$) of these families participated in a low-level intervention (Triple P level 2 or 3). (2) Two and a half years after training (between the 2012 and 2013 referent-data assessment), the site coordinator had released program material for 2785 families. (3) As presented above, we had received at least one questionnaire from each of 895 families 2 years after staff training. Taken together, practitioner self-report and material release revealed similar results while the number of completed parent-report questionnaires was smaller (43 % of the families reached based on practitioner report).

Demographic Characteristics of Samples

Demographic data for the consumer and population samples is presented in Table 1. Tests for independence indicated that population and consumer samples did not differ regarding the child's gender, $\chi^2(1, n=639)=0.45$, $p=.50$, $\phi=.03$ but did differ regarding the child's age, $t(868)=2.27$, $p=.02$, $d=.16$. Children of the population sample ($M=6.2$, $SD=3.7$) were somewhat older than those of the consumer sample ($M=5.7$, $SD=2.4$). Furthermore, as presumed, more families with low SES participated in a course compared with families from the general population: In the consumer sample, there were substantially more families in need of financial support from the state, $\chi^2(1, n=855)=61.97$, $p<.001$, $\phi=.27$ compared with the population sample. Moreover, mothers' ($\chi^2(3, n=848)=55.53$, $p<.001$, Cramer's $V=.25$) as well as fathers' educational level ($\chi^2(3, n=810)=38.62$, $p<.001$, Cramer's $V=.22$) of the consumer sample was lower compared with the population sample. Also, in the sample of participating

Table 1 Demographic statistics of the consumer and the population sample

Variable	Population sample	Consumer sample
Sample size	499	411
Age of child M (SD)	6.2 (3.7)	5.7 (2.4)
Gender of child	46.1 % girls	49.3 % girls
SES		
Delinquency of the father	4.9 %	8.8 %
Financial support from the state	8.0 %	28.3 %
Level of education (mother)		
≤10 years	12.0 %	23.7 %
11–12 years	22.7 %	33.4 %
>12 years	63.7 %	38.9 %
Other/unknown	1.6 %	4.0 %
Level of education (father)		
≤10 years	12.0 %	26.6 %
11–12 years	24.5 %	29.8 %
>12 years	59.2 %	40.7 %
Other/unknown	4.3 %	2.9 %
Migration background ^a	23.9 %	–

^a Most frequently mentioned mother countries were Russia, Kazakhstan, and Poland

families, fathers' delinquency rates were higher, $\chi^2(1, n=848)=5.17, p=.02, \phi=.08$.

Outcomes in Parenting, Child Behavior and Quality of Life

Comparison of the Consumer Sample and the Population Sample

Pre-intervention outcomes demonstrate that the consumer sample was more in need of help than the population sample (Table 2): participating parents reported more child behavior problems, SDQ $t(783)=8.75, p<.001, d=.62$, more dysfunctional parenting, PS $t(558)=7.93, p<.001, d=.66$ and a lower quality of life of their children, KINDL-R $t(783)=6.56, p<.001, d=.47$, compared with the randomly selected families of the population. In line with this result, in the consumer sample compared with the population sample, there were more children scoring in a clinical or borderline range of the SDQ, $\chi^2(1, n=785)=34.06, p<.001, \phi=.21$, and more parents showing dysfunctional parenting skills in a clinical or borderline range, PS, $\chi^2(1, n=560)=18.09, p<.001, \phi=.18$. At post-assessment, self-reported differences in negative parenting skills between population and consumer samples were non-significant, PS $t(558)=1.66, p=.12, d=.15$, indicating that means of the consumer sample were reduced to the level of the population. Nevertheless, after participating in one of the three prevention programs, parents still reported more behavior problems of their children compared with the

parents of the population sample, SDQ $t(783)=3.70, p<.001, d=.26$.

Within-Group Effects in the Consumer Sample

In a first step, paired-sample t tests were conducted to evaluate the changes in child and parenting behavior from pre to post intervention. After participation, parents reported to use less dysfunctional, PS $t(410)=-11.19, p<.001$, and more positive parenting skills, PPQ $t(410)=6.42, p<.001$. Furthermore, they rated their children to show fewer behavior problems, SDQ, $t(393)=-9.91, p<.001$, internalizing, SBQ, $t(393)=-7.36, p<.001$, conduct, SBQ, $t(393)=-8.30, p<.001$ and attention problems, SBQ, $t(393)=-8.44, p<.001$. At post-intervention, parents also reported a better health-related quality of their children's lives compared with pre-intervention, KINDL-R, $t(393)=5.63, p<.001$. Effect sizes ranged from small to moderate (Table 2).

Effect Size Comparison with Other Studies

In a second step, we compared the effect sizes from the consumer sample of the current study with those of other uncontrolled prevention trials. A meta-analysis of uncontrolled studies from a similar multi-level approach blending universal and indicated interventions revealed a pre-post within-group effect size of $ES=0.55 [0.48; 0.62]$ for parenting strategies (Nowak & Heinrichs, 2008). The effect size of the present sample was similar for dysfunctional parenting. For positive parenting, the interval of confidence of the effect in the present trial ($ES=0.34$

Table 2 Pre- and post-mean scores of outcomes in population and consumer samples

Measure	Population sample M (SD)	Consumer sample		Effect size (95 % CI)
		M (SD) Pre	M (SD) Post	
Parenting strategies				
Positive parenting (PPQ)	–	2.1 (0.4)	2.2 (0.4)	0.34 (0.20, 0.48)
Dysfunctional parenting (PS)	2.5 (0.6)	3.0 (0.8)	2.6 (0.7)	–0.56 (–0.42, –0.70)
Borderline	3.4 %	10.0 %	4.4 %	
Clinical	2.0 %	10.5 %	1.9 %	
Child behavior				
Child behavior problems (SDQ)	7.4 (5.4)	10.9 (5.8)	8.8 (5.2)	–0.50 (–0.36, –0.64)
Borderline	5.9 %	14.0 %	12.2 %	
Clinical	7.2 %	16.2 %	6.9 %	
Conduct problems (SBQ)	–	3.3 (3.9)	2.2 (3.1)	–0.41 (–0.27, –0.55)
Internalizing problems (SBQ)	–	3.2 (3.1)	2.1 (2.6)	–0.37 (–0.23, –0.51)
Attention problems (SBQ)	–	5.6 (4.1)	4.4 (3.6)	–0.42 (–0.28, –0.56)
Quality of life (KINDL-R)	80.8 (8.8)	75.9 (11.9)	79.2 (9.7)	0.28 (0.14, 0.42)

Within-group effect sizes were calculated, with $ES = (M_{\text{post}} - M_{\text{pre}}) / SD_{\text{difference}}$; the confidence interval with $CI = ES \pm z_{\alpha/2} \sqrt{s^2_{ES}}$ (Rustenbach, 2003). For participating families, sample size was $N=411$ for parenting measures and $n=394$ for child measures; for the population sample, it was $n=391$ for child behavior questionnaires and $n=149$ for the parenting measure, respectively

CI confidence interval, PPQ Positive Parenting Questionnaire, PS Parenting Scale, SDQ Strengths and Difficulties Questionnaire, SBQ Social Behavior Questionnaire

[0.20; 0.48]) lay at the lower bound of the reviewed studies. For child behavior problems, the interval of confidence of the overall effect size of the reviewed uncontrolled studies ($ES=0.57$ [0.47; 0.67], Nowak & Heinrichs, 2008) overlapped with the interval of confidence of the current study (Table 2).

Discussion

This trial is, to our knowledge, the first attempt to evaluate the community-wide implementation of evidence-based prevention programs on families reached in Germany. Our main findings demonstrate that this naturalistic population-based approach to the implementation of evidence-based prevention programs could be successfully employed in routine care settings. Moreover, many low SES families were reached and the interventions were still associated with mean positive change in families on all parameters (parenting skills, child behavior problems, quality of life). Furthermore, the delivery of programs in this context yielded effect sizes in the range of those of other uncontrolled studies (e.g., Nowak & Heinrichs, 2008). Biasing factors such as spontaneous remission were not controlled, potentially overestimating the effects. In line with this, other authors reported a loss of impact in EBP scaling. For example, in the KiVa study, the effects of an anti-bullying program were higher in a randomized controlled

study compared with a large-scale rollout (Salmivalli & Poskiparta, 2012).

Staff-evaluation results suggest that our naturalistic implementation approach was promising: 3 years after training (2013), 82 % of practitioners trained had used the program in their regular work. Thirty percent of the population participated in one of the three evidence-based interventions, although there was a significant amount of staff turnover. However, implementation needed about 1 year more for continued growth to reach the targeted 25 % of the population. Possibly, it took more time to increase a pooled demand and engage families into the interventions. Also, the practitioners may have needed some time to integrate the programs into their usual work routines. The elevated level of practitioner support following the feedback of relatively low program reach may also have contributed to the increase in program reach from 2012 to 2013. Similarly, other studies report that a lack of agent stability in practice settings (e.g., high rates of staff turnover, inadequate time to conduct courses) can lead to a delayed implementation (Mihalic, Irwin, Fagan, Ballard, & Elliott, 2004) and that full implementation usually needs 2 to 5 years (e.g., Bertram, Blase, Shern, Shea, & Fixsen, 2011). Although the full implementation needed 3 years of continued growth, presented results suggest that the families that participated needed help: All assessed parameters of socioeconomic status and parenting behavior were significantly worse in the consumer sample compared with a random sample of families from the population. This

elevated percentage of high-risk families reached in our study could be due to the less-stigmatizing way of engaging families by implementing programs into routine practice. Moreover, the participation in one of the three prevention programs was associated with positive change in families with a small to moderate within-group effect sizes.

Advantages of the Current Study

This study tested a naturalistic approach to implement EBP into routine care settings. The given findings are crucial, because they propose that EBP can be successfully disseminated in practice and reach families at risk, even when finance is limited and heterogeneity of practitioners is high. Thereby, the approach addresses several translation barriers that make EBP dissemination to real-world settings more difficult (Glasgow & Emmons, 2007) and provides a framework that—with some adaptations—could be translated to other communities. Moreover, the given approach enables the community to use the programs relatively independently of external support, which could set the foundation for the maintenance of evidence-based treatment use in family services community wide.

Several aspects of the design of the study and its implementation encourage the proposition that observed effects are probably generalizable to other communities. Facilitators who delivered courses for families were regular staff of family services or schools, mostly elementary school teachers without a college or university degree. Facilitators delivered the interventions to parents, teachers, or children at their usual workplace. Treatment fidelity was assured by delivering standardized well-evaluated trainings for facilitators, by offering regular supervision on the job and by providing free-of-cost treatment manuals as well as materials for families from the city's bureau of families.

Compared with the German population, the population sample was representative regarding financial support from the state, migration background (Statistical Federal Office 2012), child behavior problems, and quality of life (Bullinger et al., 2008; Ravens-Sieberer et al., 2008). However, there was still an overrepresentation of parents with higher education in our sample (Statistical Federal Office, 2012) restricting the generalizability of the results.

Limitations

Our findings are limited in several ways. In a real-world setting, high external validity may be achieved. On the downside of this, however, is the loss of internal validity: because of the lack of a control group, we can only examine association but no causal effects of the interventions. However, it was not the purpose of the present study to conduct an RCT but to track

changes when implementation procedures are made available in a city (in the amount possible and determined through the funding available). There is considerable face value in this design because these programs have already been shown to be efficacious in controlled studies. If these individual changes have the capacity to evoke population-level effects is another, equally important question we are currently analyzing (Frantz & Heinrichs, 2015b).

Another disadvantage of the naturalistic design is that the number of reached families could only be roughly estimated and the quality of assessed data is limited: Although material release and practitioner reports revealed similar results, we were not able to assess if one family participated in more than one intervention (e.g., child participated in EFFEKT, mother in PEP), which would overestimate the total number of families reached. Completed questionnaires were only received from a potentially biased 43 % of all families that participated, as revealed in the practitioner reports. The completed questionnaire that we received, however, included a substantial amount of missing data. Even though missing data was at random, increasing confidence in our capacity to generalize, it is unclear if the given results can be generalized to the participating families that did not answer any questionnaires. The discrepancy between estimated reach based on practitioner reports and completed questionnaires, although expected, needs to be taken into account in interpreting the findings.

Furthermore, data suggested that a small proportion of the consumer sample was also included in the population sample: 3 % of the population sample families reported that they participated in one of the three programs (2011; Frantz & Heinrichs, 2015a). This overlap (although small considering the low rate of questionnaires returned by the participating families), potentially leads to an underestimation of the differences between the samples.

We included procedures to ensure treatment fidelity (e.g., high-quality training, ongoing coaching, free-of-cost program materials for staff and consumers) and assessed the quality of staff training and the amount of program use. However, we were unable to undertake a more comprehensive assessment of intervention fidelity. Notwithstanding the above, treatment fidelity tends to be high when manualized, evidence-based programs are implemented in practice (Fagan et al., 2012; Mihalic et al., 2004). Another limitation of the present study is the focus on parent-reported data in the intervention condition only. Other outcome criteria, such as official reports on cases of child maltreatment would have strengthened the study. This data is not yet available.

Implementations Lessons

What we learn from this study is that for successful population-based EBP dissemination in the real world,

ongoing adjustment of the implementation strategies (e.g., increased practitioner support) based on staff-performance data is crucial. Therefore, it was useful that we feed-backed staff performance data to the stakeholders, site coordinator, coaches, and staff in order to enhance staff support. Also, close cooperation with key implementation partners and selecting a strong site coordinator highly improved the integration of competing philosophies and needs of programs, agencies, and scientists. Data suggested that about one third of all families participated in a low-level intervention. Accordingly, some parents seem to prefer more flexible and brief consultations. This underlines the need to provide services of different intensity to meet consumers' preferences.

Although the implementation approach used enabled most practitioners to use the program in their usual workplace, higher levels of support may further advance staff's application of the program in routine care. We strongly recommend building a more comprehensive staff-support system on site if funding permits. This includes, in addition to the applied manager's written consent to best support the new program, a further facilitation of organizational change. For example, ongoing monitoring of administrative structures (e.g., staff's case load, financial compensation) could help to identify and address potential organizational barriers to program use (Fixsen, Blase, Naoom, & Wallace, 2009). Furthermore, more frequent and compulsory individual staff supervision informed by model-pertinent data has a great potential to further enhance successful and consistent program application in the given practice setting (Fixsen et al., 2009).

In order to provide an enduring nurturing environment for all children, EBP needed to be translated with sustainability. In the present study, we tried to encourage sustainability of the implemented services by training practitioners already working in family service institutions in the city. Results demonstrate that some practitioners continued to conduct courses beyond the project period. However, for a sustainable infrastructure of community-wide, evidence-based family services, more support is needed (e.g., continued training of new practitioners, practitioner coaching, help for family recruitment; Bertram et al., 2011). In the present study, the time-limited funding of the implementation prohibited a more sustainable EBP translation. For future health promotion efforts, financial sustainability for community prevention and translational research is highly desirable to enable continuous implementation efforts.

Conclusions and Recommendations for Further Research

The present findings suggest that (a) programs developed and tested in research settings can be integrated into

routine services; (b) these programs can potentially be successfully applied by practitioners even though organizational support is limited; (c) the implementation approach used was successful in engaging hard-to-reach, high-risk families; and (d) the prevention programs can be useful when administered universally, in less controlled and more natural settings. Whether the results can be generalized to other countries or interventions is yet unclear. Studies demonstrated consistent effects for parenting programs across countries increasing confidence in the capacity to generalize (e.g., Forgatch, Patterson, & Gewirtz, 2013; Sanders et al., 2014). However, the challenge is to identify organizational change strategies that promote sustainable population-level implementation with good fidelity within diverse cultural contexts, as our service delivery systems are somewhat different. Translation research suggests that the general mechanisms are similar across countries although the social service infrastructure is different (e.g., Salmivalli & Poskiparta, 2012; Bertram et al., 2011). For example, implementation rates and practitioner dropout were comparable in European and US prevention trials (Glazemakers, 2012; Prinz et al., 2009; our study). Nevertheless, social service infrastructure influences the implementation of new programs. Potentially, it moderates the relationship between the implementation model and its outcomes. Therefore, implementation efforts should take into account the given infrastructure. Moreover, the implementation process should be continuously monitored to allow flexible use and adaptation of the implementation model using feedback loops. This adaptation of the implementation model might facilitate meeting the communities' needs and requirements when using dissemination approaches across countries. Additionally, applying similar measures and reporting conventions for international studies using the same interventions would ease replication and comprehensive use of research findings. These conclusions, as well as the implementation lessons outlined, may help other researchers and communities to carefully develop translational studies. Thereby, the study might offer another small piece in the gap of the science-to-practice puzzle in order to promote a nurturing environment for children.

Acknowledgments This research was supported by grants from the German Government of Justice [Gefördert von der Stiftung Deutsches Forum für Kriminalprävention und vom Bundesministerium der Justiz aufgrund eines Beschlusses des Deutschen Bundestages]. We are very grateful for Matthew Sanders' feedback on the first revision of this manuscript.

Conflict of Interest Nina Heinrichs and Kurt Hahlweg are members of the international scientific advisory committee for Triple P. Inga Frantz and Nina Heinrichs are members of the Triple P Research Network since 2014 (academic disclosure). Julia Plücker is co-author of PEP and conducts courses for PEP trainers (financial disclosure).

References

- Arnold, D. S., O'Leary, S. G., Wolff, L. S., & Acker, M. M. (1993). The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment*, 5, 137–144. doi:10.1037/1040-3590.5.2.137.
- Belfer, M. L. (2008). Child and adolescent mental disorders: The magnitude of the problem across the globe. *The Journal of Child Psychology and Psychiatry*, 49, 226–236.
- Bertram, R., Blase, K., Shern, D., Shea, P., & Fixsen, D. (2011). *Policy research brief: Implementation opportunities and challenges for prevention and promotion initiatives*. Alexandria, VA: National Association of State Mental Health Program Directors (NASMHPD).
- Bullinger, M., Levke Brütt, A., Erhart, M., Ravens-Sieberer, U., & the BELLA Study Group. (2008). Psychometric properties of the KINDL-R questionnaire: Results of the BELLA study. *European Child and Adolescent Psychiatry*, 17, 125–132. doi:10.1007/s00787-008-1014-z.
- Eisner, M., Nagin, D., Ribeaud, D., & Malti, T. (2012). Effects of a universal parenting program for highly adherent parents: A propensity score matching approach. *Prevention Science*, 13, 252–266. doi:10.1007/s11121-011-0266-x.
- Fagan, A. A., Hanson, K., Briney, J. S., & Hawkins, J. D. (2012). Sustaining the utilization and high quality implementation of tested and effective prevention programs using the Communities That Care prevention system. *American Journal of Community Psychology*, 49, 365–377. doi:10.1007/s10464-011-9463-9.
- Fixsen, D. L., Blase, K. A., Naoom, S. F., & Wallace, F. (2009). Core implementation components. *Research on Social Work Practice*, 19, 531–540. doi:10.1177/1049731509335549.
- Forgatch, M. S., Patterson, G. R., & Gewirtz, A. H. (2013). Looking forward the promise of widespread implementation of parent training programs. *Perspectives on Psychological Science*, 8, 682–694. doi:10.1177/1745691613503478.
- Frantz, I., & Heinrichs, N. (2014). Umgebungsgebundene Förderung: Eltern als Zielgruppe [Environmental-based promotion of child development: Parents as target group]. In A. Lohaus & M. Glüer (Eds.), *Entwicklungsförderung im Kindesalter*. Göttingen, Germany: Hogrefe.
- Frantz, I., & Heinrichs, N. (2015a). Effekte einer Werbekampagne für ausgewählte evidenzbasierte Präventionsprogramme [Effects of a communication strategy for evidence-based prevention programs]. *Kindheit und Entwicklung*, 24, 20–27.
- Frantz, I. & Heinrichs, N. (2015b). *Populationseffekte einer flächendeckenden Implementierung familienbasierter Präventionsprogramme [Population-level effects of family-based prevention programs]*. Manuscript in preparation.
- Glasgow, R. E., & Emmons, K. M. (2007). How can we increase translation of research into practice? Types of evidence needed. *Annual Review of Public Health*, 28, 413–433. doi:10.1146/annurev.publhealth.28.021406.144145.
- Glazemakers, I. (2012). *A population health approach to parenting support: Disseminating the Triple P- Positive Parenting Program in the Province of Antwerp*. Dissertation, University of Antwerp, Antwerp.
- Goodman, R. (1997). The strengths and difficulties questionnaire: A research note. *Journal of Child Psychology and Psychiatry*, 38, 581–586. doi:10.1111/j.1469-7610.1997.tb01545.x.
- Lösel, F., & Stemmler, M. (2012). Preventing child behavior problems in the Erlangen-Nuremberg Development and Prevention Study: Results from preschool to secondary school age. *International Journal of Conflict and Violence*, 6, 214–224.
- McMahon, R. J., & Forehand, R. L. (2003). *Helping the noncompliant child: Family-based treatment for oppositional behavior*. New York: Guilford Press.
- Mihalic, S., Irwin, K., Fagan, A., Ballard, D., & Elliott, D. (2004). Successful program implementation: Lessons from blueprints. *Juvenile Justice Bulletin*. Retrieved May, 13, 2014, from <https://www.ncjrs.gov/pdffiles1/ojdp/204273.pdf>.
- National Research Council & Institute of Medicine. (2009). *Preventing mental, emotional, and behavioral disorders among young people: Progress and possibilities*. Washington, DC: National Academies Press.
- Naumann, S., Bertram, H., Kuschel, A., Heinrichs, N., Hahlweg, K., & Döpfner, M. (2010). Der Erziehungsfragebogen (EFB)-Ein Fragebogen zur Erfassung elterlicher Verhaltenstendenzen in schwierigen Erziehungssituationen [The Parenting Questionnaire. An instrument to assess parental behavioral tendencies in discipline situations with their children]. *Diagnostica*, 56, 144–157. doi:10.1026/0012-1924/a000018.
- Nowak, C., & Heinrichs, N. (2008). A comprehensive meta-analysis of Triple P-Positive Parenting Program using hierarchical linear modeling: Effectiveness and moderating variables. *Clinical Child and Family Psychological Review*, 11, 114–144. doi:10.1007/s10567-008-0033-0.
- Olds, D. L. (2000). The nurse-family partnership: An evidence-based preventive intervention. *Infant Mental Health Journal*, 27, 5–25. doi:10.1002/imhj.20077.
- Plueck, J., Eichelberger, I., Hautmann, C., Hanisch, C., Jaenen, N., & Doepfner, M. (2014). Effectiveness of a teacher-based indicated prevention program for pre-school children with externalizing problem behaviour. *Prevention Science*, 16, 233–241. doi:10.1007/s11121-014-0487-x.
- Prinz, R. J., Sanders, M. R., Shapiro, C. J., Whitaker, D. J., & Lutzker, J. R. (2009). Population-based prevention of child maltreatment: The U.S. Triple P system population trial. *Prevention Science*, 10, 1–12. doi:10.1007/s11121-009-0123-3.
- Ravens-Sieberer, U., Wille, N., Erhart, M., Bettge, S., Wittchen, H.-U., Rothenberger, A., ... as the BELLA study group (2008). Prevalence of mental health problems among children and adolescents in Germany: Results of the BELLA study within the National Health Interview and Examination Survey. *European Child Adolescent Psychiatry*, 17, 22–33. doi:10.1007/s00787-008-1003-2.
- Rustenbach, S. J. (2003). *Metaanalyse. Eine anwendungsorientierte Einführung [Meta-analysis: A user-oriented introduction]*. Bern, Switzerland: Huber.
- Salmivalli, C., & Poskiparta, E. (2012). KiVa antibullying program: Overview of evaluation studies based on a randomized controlled trial and national rollout in Finland. *International Journal of Conflict and Violence*, 6, 294–302.
- Sanders, R. M., Kirby, J. N., Tellegan, C. L., & Day, J. J. (2014). The Triple P-Positive Parenting Program: A systematic review and meta-analysis of a multi-level system of parenting support. *Clinical Psychological Review*, 34, 337–357.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7, 147–177.
- Shure, M. B. (1992). *I Can Problem Solve: An interpersonal cognitive problem-solving program*. Champaign: Research Press.
- Spoth, R., Rohrbach, L. A., Greenberg, M., Leaf, P., Brown, C. H., Fagan, A., ... Hawkins, J. D. (2013). Addressing core challenges for the next generation of type 2 translation research and systems: The translation science to population impact (TSci Impact) framework. *Prevention Science*, 14, 319–351. doi:10.1007/s11121-012-0362-6.
- Statistical Federal Office (2012). *Statistical Yearbook for the Federal Republic of Germany*. Statistical Federal Office. Retrieved May, 15, 2014, from <https://www.destatis.de/DE/Publikationen/StatistischesJahrbuch/StatistischesJahrbuch2012.pdf>.
- Strayhorn, J. M., & Weidman, C. S. (1988). A Parent Practices Scale and its relation to parent and child mental health. *Journal of the American Academy of Child and Adolescent Psychiatry*, 27, 613–618.

- Tremblay, R. E., Vitaro, F., Gagnon, C., Piché, C., & Royer, N. (1992). A prosocial scale for the Preschool Behavior Questionnaire: Concurrent and predictive correlates. *International Journal of Behavioral Development*, 15, 227–245. doi:[10.1177/016502549201500204](https://doi.org/10.1177/016502549201500204).
- Welsh, B. C., & Farrington, D. P. (2012). *The Oxford handbook of crime prevention*. Oxford: University Press.