Journal of Consulting and Clinical Psychology
2015, Vol. 83, No. 6, 1069–1084

0022-006X/15/$12.00 http://dx.doi.org/10.1037/a0039328

A National Evaluation of Parenting Programs in Sweden: The Short-Term Effects Using an RCT Effectiveness Design

Håkan Stattin
Örebro University

Pia Enebrink
Karolinska Institutet

Metin Özdemir
Örebro University

Fabrizia Giannotta
Uppsala University

Objective: We evaluated the effectiveness of 4 parent-training programs for children with externalizing problems. We tested the effectiveness of 3 behavioral programs (Comet, Cope, and Incredible Years) and 1 nonbehavioral program (Connect) in reducing child behavior problems and attention-deficit/hyperactivity disorder (ADHD) symptoms, in improving positive parenting and parenting competence, and in decreasing negative parenting and parents' stress and depressive symptoms. ***Method:*** This national study was designed as a randomized-controlled effectiveness trial (RCT). The treatments were carried out in 30 clinical and community-based practices. Parents of 908 children (ages 3–12 years) were randomly assigned to 1 of 2 parent training programs available at each practice, or to a wait-list condition, where parents had sought help from regular services. Before and after treatment, parents rated child behavior problems and parenting strategies. ***Results:*** At posttreatment, children whose parents had received interventions showed a strong decrease in child conduct problems and a moderate to strong decrease in ADHD symptoms. About half of parents whose children scored over the 95th percentile on the behavior measures (Eyberg Child Behavior Inventory, Swanson, Nolan, and Pelham Rating Scale), a clinically meaningful cutoff, reported that their children were no longer above the cutoff after the intervention. Parents showed considerably less negative behaviors toward their children at posttest compared with pretest; they increased in parental competence, and decreased in both stress and depressive symptoms. Overall, the behavioral programs were more effective than the nonbehavioral program. ***Conclusions:*** The results support the general efficacy of parent training in a short-term perspective.

What is the public health significance of this article?
This effectiveness evaluation examined 4 group-based parenting programs with different theoretical foundations. All 4 of the programs decreased child problem behaviors substantially and increased parents' sense of competence.

Keywords: parent training program, effectiveness trial, conduct problems, ADHD, children

Over the past four decades, the primary procedure for helping parents who have considerable problems in managing their children's externalizing problems has shifted from individual therapy in a clinic to outpatient parenting activities in a group setting. These programs are often structured, typically involve role-playing or video recordings to teach effective parenting skills, and encourage reflection and practice over three to four months. Do these parenting programs, as used in selective prevention, help parents and children with externalizing problems?

A large number of published studies have attempted to address this question. A number of meta-analyses have covered both behavioral and nonbehavioral programs (Cedar & Levant, 1990; Lundahl, Risser, & Lovejoy, 2005; Maughan, Christiansen, Jenson, Olympia, & Clark, 2005; Serketich & Dumas, 1996), and some 20 review studies (e.g., Farmer, Compton, Burns, & Robertson, 2002; Nixon, 2002) have tried to summarize the results. The consensus is that there are improvements in child behavior and the programs generally help parents in the short term, with effect

This article was published Online First May 25, 2015.

Håkan Stattin, Youth&Society, School of Law, Psychology, and Social Work, Örebro University; Pia Enebrink, Department of Clinical Neuroscience, Division of Psychology, Karolinska Institutet; Metin Özdemir, Youth&Society, School of Law, Psychology, and Social Work, Örebro University; Fabrizia Giannotta, Department of Psychology, Uppsala University.

This research made possible by a grant from The National Board of Health and Welfare. We thank all participating parents, children, and professionals who implemented the parenting programs.

Correspondence concerning this article should be addressed to Håkan Stattin, Youth&Society, School of Law, Psychology, and Social Work, Örebro University, SE-701 82 Örebro, Sweden. E-mail: hakan.stattin@oru.se

sizes (Cohen's *d*) ranging from around .30 to .40 (Litschge, Vaughn, & McCrea, 2010; Maughan et al., 2005; Serketich & Dumas, 1996). However, meta-analyses and reviews often gather together studies that differ in their degree of methodological rigor. For example, they mix randomized controlled trials and quasi-experimental studies, and programs addressing parenting only and ones with multiple components, which makes it difficult to draw firm conclusions about the effectiveness of particular types of programs. In an attempt to overcome this limitation, Furlong et al. (2012) applied more stringent criteria and restricted their selection of studies only to randomized controlled trials (RCTs) of parenting programs designed to prevent children's externalizing problems. By so doing, they only found 13 evaluations of 5 programs, with interventions covering a total of 1,078 parents of children in the ages 3–12 years. The authors reported moderately favorable short-term effects of the programs on child disruptive behaviors, parental mental health, positive parenting, and reductions in harsh, inconsistent parenting—results that were replicated in both parent reports and independent evaluations.

One question that remains unanswered concerns the types of programs that are most effective for parents and children. Comparisons in meta-analyses are based on studies that differ in study-population size, settings, design, instruments, and data quality. Thus, their conclusions may not be robust. There is currently only one empirical study in the literature that has systematically analyzed differences between the programs in one and the same study. Lindsay and Strand (2013) evaluated four programs for parents—Incredible Years, Strengthening Families 10–14, Strengthening Families Strengthening Children, and Triple P—of children 8–13 years. They reported moderate to strong effect sizes at posttest for child conduct problems and measures of parenting and parent mental well being, and few differences between programs. These effects were sustained at the 1-year follow-up. The robustness of their conclusions is limited, though, because almost half of parents were lost at posttest, and only one in six of parents with data at baseline reported at the 1-year follow-up. In addition, the implementations were not conducted as part of normal regular practices, and no control group was used. This leaves open the general question of scientific, economic, and public interest: Which programs are best at producing clinically relevant results?

Another limitation to current research lies in a lack of effectiveness trials. Most of the studies reported in the literature are efficacy trials conducted at a research institution, and there are few clinically based randomized controlled effectiveness trials of parent management training (PMT) programs (see Costin & Chambers, 2007 an exception; Kling, Forster, Sundell, & Melin, 2010; van den Hoofdakker et al., 2007). Effectiveness studies are important for answering the question of whether any program effects detected in efficacy studies can also be observed in real-life clinical settings. Some very successful parenting programs, such as Triple P, failed to show effectiveness when implemented in such settings, whereas others, such as Incredible Years, retained their effects (Little et al., 2012). On the other hand, a recent meta-analysis (Michelson, Davenport, Dretzke, Barlow, & Day, 2013) contrasting results from eight parenting programs evaluated across 28 RCTs did not find any significantly lower effect sizes for studies that used strict real-world criteria (close to being effectiveness trials) than those applying less stringent real-world criteria. Given the current official recommendation to implement parenting programs in clinical contexts, these findings are encouraging (NICE, 2006). However, differences in the observed outcomes make it difficult for clinicians, researchers, and policymakers to understand what works under what circumstances, and prohibits the drawing of firm conclusions about the effects of parenting programs. Thus, more effectiveness trials are needed to understand the impact of parenting programs in real-life settings.

The Present Study

In the present study, we tried to overcome some of the limitations of earlier research. In an attempt to understand which programs are most effective in real-life settings, we evaluated the most common parenting programs employed in Sweden, and which are delivered as usual in clinical and community service settings. The most widespread programs when the study started were: *Comet* (Kling et al., 2010), *Cope* (Cunningham, 2005), *Incredible Years* (Webster-Stratton, Reid, & Hammond, 2004), and *Connect* (Moretti, Holland, Moore, & McKay, 2004). The first three programs are all inspired by social learning theory, and aim to strengthen children's social skills and emotional regulation by using praise and incentives to encourage cooperative behavior, by ignoring children's inappropriate behaviors, and by exerting positive discipline through rules, routines and effective limit setting. *Comet* is a Swedish program inspired by Patterson, Dishion, & Chamberlain (1993) Parent Management Training–Oregon Model. Comet also contains an individual consultation session in addition to having parent groups. The program is probably the most culturally appropriate in the context of the four programs in Sweden, because it was initially developed for Swedish families. *Incredible Years* is built on Patterson's coercion model, Bandura's (1986) notions of modeling and self-efficacy, Piaget & Inhelder (1962) developmental interactive learning methods, and cognitive strategies for challenging angry, negative and depressive self-talk. *Cope* is different from the two behavioral programs just described to the extent that parents are supposed to work together and generate their own solutions to problems. Thus, in spite of their common roots in the behavioral approach, the three programs have their own peculiarities, which can differentially affect parents and children's behaviors (see Table 1 for a description of the Swedish versions of the programs).

In contrast to the three behavior-based programs, *Connect* is an attachment-focused program that encourages parents to reflect on the parent–child relational context, and the attachment needs of the child (Moretti & Obsuth, 2009; Obsuth, Moretti, Holland, Braber, & Cross, 2006). It strives to stimulate curiosity and reflection on the basis of information on various attachment themes, rather than giving parents specific skills in behaving as a parent. The program focuses on teaching parents how to reframe child behaviors, how to change their own emotional responses when the child acts inappropriately, and how to communicate empathy, all of which are achieved through reflexive exercises, role-plays, and discussions. Thus, there are similarities and dissimilarities between the behavioral programs and the nonbehavioral program. Similarities include a focus on parenting as a context, and also on the use of reflections, discussions, and role-plays, but Connect (the nonbehavioral program) differs from the behavior-based programs in lacking homework and having a sharper focus on relation-oriented child needs.

Table 1
The Features of the Swedish Versions of the Parent Training Programs: Comet, Cope, Incredible Years, and Connect

Feature	Comet	Cope	Incredible Years	Connect
Aims	Originally developed for parents to teach how to manage negative child behaviors; also being used for children diagnosed with ADHD, ADD, and ODD	Originally developed for parents of pre-adolescents diagnosed with ADHD; also being used for the parents of children with ADHD, ODD, CD, parents of disadvantaged background (e.g., minorities)	Originally developed as an intervention for children with ODD; also being used for parents of high-risk, and children with behavioral problems	Originally developed for parents of preteens and teens who struggle with significant behavioral problems (CD, aggressiveness, violence, antisocial behavior, delinquency) and other mental health issues (concurrent anxiety and depression, substance use problems)
Theory	Developed based on Webster-Stratton's, and Patterson's (1982) and Barkley's (1997) parent management models: cognitive behavioral therapy	Social learning theory; also includes principles of cognitive and social psychological models on attitude change, family systems theory, and small-group interventions	Cognitive social learning theory, particularly Patterson's (1982) coercion model; Bandura's (1986) notions of modeling and self-efficacy; Piaget & Inhelder (1962) developmental interactive learning methods	Attachment theory; systemic theories; relational theories
Age range	3–12 years	3–12 years	3–8 years	9–16 years (in this study the age range was 8–12 years)
Sessions	11 2.5-hr weekly sessions 10–12 parents (6 families)	10 1-hour weekly sessions Maximum 25 parents	12 2.5-hr weekly sessions 10–14 parents	10 1-hour weekly sessions 12–14 parents
Modalities	Teaching Video vignettes Role plays Homework Take-home materials for parents One individual meeting	Modeling Group discussions Role plays Homework Self-monitoring	Teaching Group discussions Videotaped modeling Role plays and rehearsal Weekly homework, evaluations Phone calls, make-up sessions Buddy calls	Teaching Role playing Take-home handouts for parents

Note. All programs were delivered by two group leaders, and implemented in child–adolescent psychiatry clinics, social services, and schools. ADHD = attention-deficit/hyperactivity disorder; ADD = attention deficit disorder; ODD = oppositional defiant disorder; CD = conduct disorder.

All of the four parenting programs that are included in the current study have been evaluated, and found effective on a number of child and parent outcomes. First, a randomized effectiveness trial of Comet reported large effects of program on children's conduct problems and moderate effects on parents' competence compared with wait-list condition at posttest (Kling et al., 2010). The effect sizes for conduct problems were moderate with Cohen's *d* values ranging between .30 and .46. Second, Incredible Years program has been widely evaluated by both the program developer (e.g., Menting, Orobio de Castro, & Matthys, 2013; Reid, Webster-Stratton, & Hammond, 2003; Webster-Stratton, Hollinsworth, & Kolpacoff, 1989) and independent researchers (Axberg & Broberg, 2012). The Swedish evaluation (Axberg & Broberg, 2012) reported large improvements in disruptive behaviors among the intervention group compared with a wait-list condition, and the improvements were retained over one year. The program effect on mothers' stress was small whereas the effects on parental control and perceived parent–child alliance were small to moderate. Furlong et al. (2012) summarized all published effect sizes for conduct problems, and reported standardized mean difference effect sizes ranging between .24 and 1.29 across studies. Next, the comparison of the Community Parent Education (COPE) program against a wait-list control and a clinic-based individual parent training also provided evidence for the effectiveness of the program (Cunningham, Bremner, & Boyle, 1995). COPE participants were found to show greater improvements in child behavior problems compared with comparison conditions. In addition, these effects were maintained over a six-month period. Thorell (2009) reported Cohen's *d* effect sizes for COPE ranging between .24 and .94 across clinical and community sample. Finally, the attachment based Connect program was found effective in both small and larger scale efficacy trials (pre- to postevaluation and a 12-month follow-up) of parents in community settings participating in Connect compared with a wait-list control period (Moretti & Obsuth, 2009). The program was found effective in reducing youth externalizing and internalizing problems, rule-breaking, and aggressive behaviors. The effect size of the program on externalizing problem was moderate (Cohen's $d = .46$). Parents who received the Connect program also displayed improvements in satisfaction and parenting efficacy. Overall, these four parent-training programs showed significant improvements in both child and parents outcomes compared with a no-treatment wait-list condition. Nevertheless, the lack of consistency in measured outcomes and design differences make comparison of the effectiveness of these programs relative to each other make cumbersome.

What distinguishes this study from previous studies in the literature is that several parenting programs are evaluated simultaneously. We used a three-group RCT design, in which parents were randomized to one of at least two program groups or a wait-list control group, and where the same instruments were used

Table 2
Means and Standard Deviations of the Child and Parent Outcomes Across Programs and the Control Group

	Comet				Cope				Incredible Years				Connect				Waitlist control			
	Pretest		Posttest		Pretest		Posttest		Pretest		Posttest		Pretest		Posttest		Pretest		Posttest	
Child outcome	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
ECBI Intensity	3.6	0.9	2.7	0.8	3.6	0.8	2.8	0.8	3.7	0.9	3.0	0.8	3.6	1.0	2.8	1.0	3.5	0.9	3.2	0.9
≥95%	39.0	48.9	8.5	28.0	42.6	49.6	11.2	31.6	50.4	50.2	11.0	31.4	40.3	49.2	16.5	37.2	31.7	46.7	17.8	38.4
ECBI Problems	0.4	0.2	0.2	0.2	0.4	0.2	0.3	0.2	0.5	0.2	0.3	0.2	0.4	0.3	0.3	0.2	0.4	0.3	0.3	0.2
≥95%	52.0	50.1	18.5	38.9	52.5	50.1	30.2	46.1	62.5	48.6	32.7	47.1	57.1	49.6	31.8	46.7	50.3	50.2	40.4	49.2
SNAP-IV Inattention	1.1	0.8	0.8	0.6	1.1	0.7	0.9	0.7	1.2	0.8	0.9	0.7	1.3	0.7	1.0	0.7	1.0	0.7	0.9	0.7
≥95%	19.0	39.4	7.9	27.0	20.9	40.8	12.8	33.5	20.3	40.4	10.9	31.3	26.3	44.1	15.4	36.2	12.1	32.7	9.5	29.5
SNAP-IV Hyperactivity	1.1	0.8	0.8	0.7	1.2	0.7	0.9	0.7	1.3	0.8	0.9	0.6	1.0	0.8	0.8	0.7	1.1	0.8	1.0	0.7
≥95%	30.7	46.3	18.2	38.7	36.6	48.3	23.9	42.8	43.8	49.8	23.9	42.8	34.0	47.5	19.8	39.9	32.5	47.0	28.1	45.1
SNAP-IV ODD	1.2	0.7	0.8	0.6	1.1	0.6	0.8	0.6	1.2	0.7	0.8	0.6	1.2	0.8	0.9	0.7	1.1	0.7	0.9	0.6
≥95%	14.5	35.3	7.3	26.2	7.9	27.1	5.3	22.5	15.6	36.4	6.4	24.5	20.0	40.1	12.0	32.6	11.4	31.9	8.2	27.6
Negative parenting																				
Angry Outbursts	2.1	0.5	1.8	0.4	2.1	0.5	1.8	0.4	2.0	0.5	1.9	0.5	2.0	0.5	1.9	0.5	2.1	0.5	1.9	0.5
Harsh Parenting	2.3	0.6	1.8	0.4	2.3	0.7	1.9	0.5	2.3	0.7	2.0	0.5	2.3	0.7	2.0	0.6	2.4	0.7	2.2	0.6
Positive parenting																				
Attempted understanding	2.6	0.4	2.7	0.3	2.6	0.4	2.7	0.3	2.6	0.3	2.7	0.4	2.6	0.3	2.7	0.3	2.6	0.3	2.6	0.3
Rewards	3.1	1.2	3.4	1.2	3.4	1.5	3.5	1.3	3.5	1.3	3.6	1.3	3.3	1.3	3.5	1.5	3.3	1.3	3.1	1.3
Parenting competence																				
PSOC	3.9	0.7	4.5	0.6	3.8	0.7	4.3	0.7	3.9	0.7	4.2	0.7	3.7	0.6	4.2	0.7	3.8	0.6	4.0	0.6
Parental mental health																				
Stress	2.1	0.7	1.6	0.5	2.1	0.8	1.8	0.7	2.2	0.8	1.8	0.6	2.2	0.8	1.8	0.7	2.1	0.7	1.8	0.7
Depression	0.9	0.6	0.7	0.6	1.0	0.6	0.8	0.6	0.9	0.6	0.8	0.6	1.0	0.6	0.8	0.6	0.9	0.6	0.9	0.6
≥95%	29.6	45.8	18.6	39.1	34.7	47.7	23.3	42.4	31.2	46.5	27.3	44.7	38.1	48.7	23.4	42.4	32.1	46.8	31.8	46.7

Note. ECBI = Eyberg Child Behavior Inventory (Colvin, Eyberg, & Adams, 1999); SNAP-IV = Swanson, Nolan, and Pelham Rating Scale (Swanson, Nolan, & Pelham, 1992); ODD = oppositional defiant disorder; PSOC = Parents' Sense of Competence measure (Johnston & Mash, 1989).

to compare outcomes between the groups. This procedure should represent an improvement from relying on effect sizes from meta-analyses of studies that differ considerably in design and data quality. In this study, we asked whether some of the commonly used parenting programs in Sweden are equally effective in reducing disruptive child behaviors, promoting good parent–child interactions, and increasing parents' perceptions of their parenting competence and psychological well-being. Because this is the first time multiple programs with different theoretical backgrounds are compared with each other in one and the same study, specific hypotheses are not stated.

Our study was designed to evaluate the effectiveness of the parenting programs in regular practice. To achieve this goal, first, we did not change anything in the usual practices of the service units, except for holding information meetings with the parents where they were informed about the study and asked for consent. Second, our exclusion criteria were lax. We included all children who were between the ages of 3 and 12 years without an autism spectrum disorder diagnosis. Third, we purposefully allowed parents in the intervention and wait-list conditions to seek help from other available services so that we could compare the effectiveness of the parenting programs with that of what is counted as treatment as usual in Sweden.

We also wanted to evaluate the clinical relevance of the program outcomes by examining how many children who were above the 95th percentile cutoff for children in normal samples at pretest showed improvements by the end of the parenting programs. We further examined whether the four parenting programs effected similar or dissimilar changes in parents' negative and positive parenting practices, parenting competence, and parental stress and depressive symptoms. Finally, we examined possible moderators. Because some studies have shown moderating effects of sociodemographic characteristics on the effectiveness of PMT programs (Lundahl et al., 2005; Reyno & McGrath, 2006), we examined the effects of a number of potential moderating factors, such as child age and gender, parent age, family income, economic (personal financial) strain, marital status, immigrant origin, and child use of concomitant medication.

Method

Design and Study Setting

The study was designed as a randomized-controlled effectiveness trial, following the criteria laid out in Gartlehner et al. (2006) for conducting an effectiveness study. Four research groups at the Örebro University, Karolinska Institute, Göteborg University, and Lund University representing four of the most populated administrative regions of Sweden, collaborated on the research. In each administrative region, the human services units (e.g., schools, social welfare agencies, and child and adolescent psychiatry clinics) that had implemented at least two of the targeted parenting programs were identified. Overall, we found 30 different units that had implemented at least two of the four parenting programs. All these 30 units are included in the study.

All the units were already running the programs with their regular personnel, and they did not change their routine practices when they became part of the evaluation. The primary care units recruited parents in accordance with their usual protocols. Most parents had contacted a unit on their own, but a few were recruited

through advertisements about the availability of parenting programs in their communities (which was also a part of normal routine in these communities). When large enough groups of interested parents had been formed, the parents were invited to an information meeting held by the research personnel. At the meetings, parents were told that a scientific evaluation of different parenting programs was underway, and that they were eligible for recruitment into the study. The arrangement of these information meetings was the only change to the usual practice of the units brought about by the current research program. All parents gave active consent to their participation in the study. After the meetings, our research personnel sent out the results of the randomization to the parents (in some cases, research assistants provided parents with the results of randomization at the meeting, after informed consent had been given, and the questionnaire was filled in). Then, the parenting programs began. In total, there were 72 separate parenting groups. As part of the usual practice of the units, interpreters were engaged in a few cases to assist parents of immigrant origin. Pretest, in this study, is the time when the parents responded to our first questionnaire. This was at the information meeting, one to two weeks before the parenting program started. Posttest is three to four months after pretest, when the parents had completed their parenting program. The second questionnaire was sent by regular mail to the homes of the parents. We define short-term effects in terms of changes that took place between pretest and posttest.

Randomization

We randomized parents into three groups at each treatment unit after the number of parents interested in participating had been estimated. Because the Ethics Committee did not permit the inclusion of a control condition in the first year of the study, we included in the design a condition where parents were given a book on parenting management and instructions on how to use the book (i.e., a self-help-book condition). We were allowed to include a wait-list control condition in the second year. Thus, in the first year of the study, the parents were randomly assigned to one of three conditions: (a) one of the programs available at the unit, (b) the other program offered on the site, and (c) a self-help-book condition. The book condition is not considered in the current paper, and will be reported separately. For the second year, the parents were randomly assigned to one of the following conditions: (a) one of the programs available at the unit, (b) the other program offered on the site, and (c) a wait-list condition.

In the present study, we compare all parents who, during the two years, were randomized to a parenting program, with the parents who, during the second year, were randomized to the wait-list. Only including the parents who, during the second year, were randomized to a parenting program generated virtually the same results as including all the parents who, during both the first and the second year, were randomized to a program. Thus, in order to increase statistical power, we chose to include all parents who were randomized to one of the four parenting programs during the two years.

Age was taken into account in the randomization, because the preadolescent version of the Connect program was designed for children older than 8 years old, whereas the Incredible Years program was aimed at younger children. For children from 3 to 8 years, parents were randomized to the age-relevant versions of Cope, Incredible Years, and Comet. For children from 9 to 12 years, parents were randomized to the programs that were developed for older children, namely Cope, Comet, and Connect.

Participants

The parents of 1113 children attended the first information meetings. Three children did not meet the inclusion criteria, and were excluded prior to randomization: one child younger than 3 years, one older than 13 years, and one diagnosed with an autism spectrum disorder. Also, the service units determined that 6 additional children would not benefit from manual-based parenting programs, and were therefore excluded from the study groups. Overall, parents of 1,104 children were randomized according to the study conditions. The sample of the present study is composed of the parents and children randomized to one of four parenting programs or the wait-list condition. Those randomized to the self-help-book condition ($n = 196$) were not included in the current evaluation. Hence, the final sample comprised 908 children and their parents: 749 assigned to one of the programs, and 159 assigned to the wait-list condition. Descriptions of enrolment, sample allocations and dropout are presented in Figure 1.

At pretest, 24.6% of the parents had extra continuous help at school or at home (from a special education teacher, counselor, psychologist, school psychologist, assistant, speech therapist, social worker, contact person, etc.), and there were no significant differences between the parents on a program, or between the parents in the wait-list condition. At posttest, 21.4% of the parents said that they received support. Again, there were no significant differences between the parents on the separate programs and the parents in the wait-list condition (fewest, 17.6%, for the wait-list parents; most, 24.5%, for parents on the Connect program).

At baseline, the most commonly reported diagnosis among the children was attention-deficity/hyperactivity disorder (ADHD) (48%) and allergies and/or asthma (25%). No other diagnosis comprised more than 3%. Distribution of children with an ADHD diagnosis did not significantly differ across the programs, $\chi^2(3) = 6.36$, $p = .095$. Of children with ADHD ($N = 93$; 8%), 60% were treated with stimulant drugs including methylphenidate (Ritalin, Concerta, Equasym Depot, Medikinet) or amphetamines, and 12% with nonstimulant drugs (Strattera). The use of these mediations did not significantly differ across the programs. However, use of concomitant medication was more common among children whose parents participated in the programs at child and adolescent psychiatry clinics (44%) and social welfare (40%) compared with schools (6.3%) and primary care (9.7%).

The main source of attrition in the study was not having started on a parenting program. Of the parents of the 749 children who were randomized to a parenting program, 635 (84.8%) subsequently started on a program. Of these 635 parents, we had complete pre- and posttest data on 598 (79.8% of those who were originally randomized to a parenting program; 94.2% of those who had started on a program). Fewer parents started on the Incredible Years program (75.4%) than on the other programs. The lower frequency of parents on the Incredible Years program was due to organizational problems. In two of the communities where Incred-

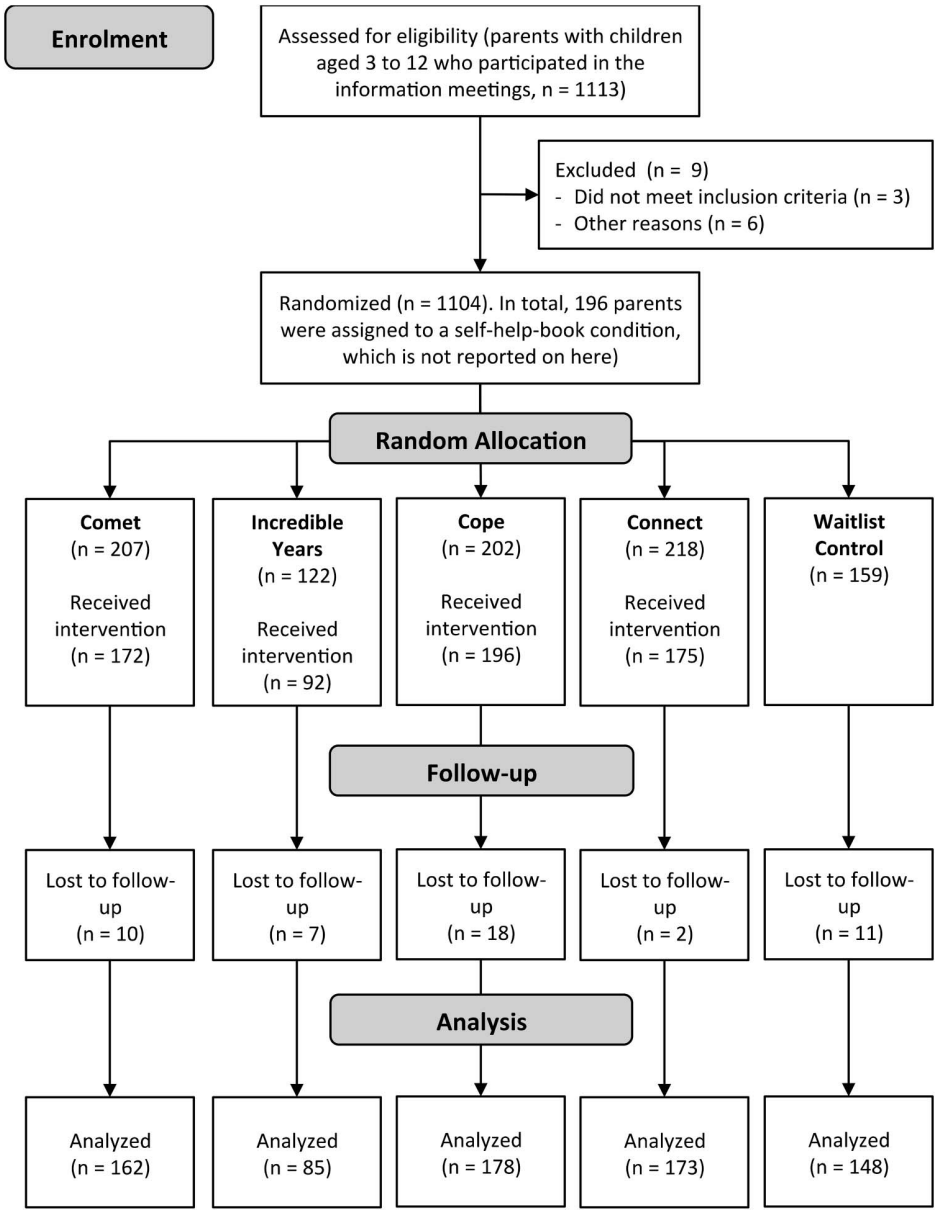

Figure 1. Consolidated Standards of Reporting Trials (CONSORT) diagram: sample randomization, program participation, and assessments.

ible Years was launched, local communities had to collaborate with other communities to recruit enough parents. Because of the geographical location of the sites, most of the parents recruited in neighboring communities had to travel long distances to take part in the program, and as a result, many chose not to attend. However, once parents had decided to start, a large majority (93.6%) maintained their participation and provided data at both pre- and posttest. Of the 159 parents in the wait-list condition, posttest data were available for 148 parents (93.1%). Regarding the statistical power of the study, the sample size was adequate for detecting small effects. All programs had the recommended sample size of at least 75 parents per program (Gartlehner et al., 2006). Indeed, the number of participants was more than twice that required for Cope, Comet, and Connect. Thus, the power of the study should be considered satisfactory.

In most cases, one parent in the family (78.1%), usually the mother but in some cases the biological father, attended the program meeting. In about a fifth of cases (21.9%), both parents attended the parent training sessions. For the present study, we selected the parent who had participated in most sessions of the parenting groups as the primary reporter of the child's problem behaviors and other aspects of the parenting situation. If the number of attendances was equal between parents, we chose the mother. Overall, mothers were the primary reporters among both

those who attended the programs (85.0%) and those in the wait-list condition (86.1%). Note that the total 1,104 parents recruited to the study were the primary reporters.

In two out of three cases, the child targeted by the programs was a boy. The average age of the child was 7.70 years (SD = 2.60), and ranged between 3 and 12 years. The average age of the primary reporter was 37.7 years (SD = 7.51), with the youngest age 20, and the oldest, 60. About three out of four were married or cohabiting (74%), and the rest were single parents. In most cases (89%), both parents were born in one of the Scandinavian countries. The average monthly household income after tax was 30,000 to 40,000 SEK ($3,500 to $4,700). There were 6.1% whose monthly incomes were as low as 0–10,000 SEK ($0-$1,200), and 24.9% had an income higher than 50,000 SEK ($5,900). A majority of the reporters (62%) reported that their monthly income was adequate for their household expenses, but 6.3% acknowledged that their monthly income was not fully adequate. Finally, 45.5% of the parents had completed some university-level education, and 9% had only a compulsory-school education.

Attrition Analyses

At baseline, we had complete data on all the parents who were randomized to any of the conditions, except for three families. We examined whether the baseline characteristics of the parents and children were related to starting on a parent-training program by using logistic regression analysis. In the logistic regression model, starting-on-a-parenting-program was the dependent variable (1 = *did not start*; 0 = *started*), and all of the child and parent outcomes and the demographic characteristics of the children and families were entered as predictors. The results showed that none of the study variables was significantly related to parents' deciding to start on a program.

We used another logistic regression model to test whether the parents lost to follow-up in the posttest measurements (0 = *participated*; 1 = *lost to follow-up*) were predicted by baseline characteristics. The results suggested that only two of the outcome variables and one of the demographic characteristics were related to nonparticipation in posttest assessment. Specifically, parents with an immigrant background (odds ration [OR] = 2.26, p = .026), and those with higher scores on the Parents' Sense of Competence measure (PSOC; Johnston & Mash, 1989) measure (OR = 1.49, p = .040) and depressive symptoms (OR = 1.92, p = .002) were more likely not to participate in posttest assessment. Overall, given a low Nagelkerke R^2 (= .07), we concluded that attrition in the data would only lead to minimal bias in the study results.

Measures

At pre- and postassessments, parents responded to statements about child problem behaviors, ADHD symptoms, negative and positive reactions to the child, perceptions of their parenting competence and their emotional well-being, and also to demographic questions about themselves and their child.

Child Behavior Measures

Externalizing problems. We assessed children's problem behaviors, and attention deficit and hyperactivity symptoms. To assess externalizing problems, we used the Eyberg Child Behavior Inventory (ECBI; Colvin, Eyberg, & Adams, 1999). The ECBI is a widely used 36-item inventory of child problem behaviors that has both intensity and problems subscales. On the Intensity subscale, parents rate the intensity of their children's problems on a 7-point scale, indicating the frequency of each behavior (α = .92 and 94 at pre- and posttest, respectively). The problem subscale elicits *yes–no* responses to whether parents perceive their child's behavior as problematic (α = .92 on both occasions). We used Swedish norm values (Axberg, Johansson Hanse, & Broberg, 2008) to examine the clinical relevance of the intervention effect, comparing the proportions of children at pre- and posttest who were above the level of the 95th percentile of a normal sample of Swedish children.

Attention deficit and hyperactivity problems. ADHD ratings based on the *Diagnostic and Statistical Manual of Mental Disorders* (4th ed.; American Psychiatric Association, 1994) were measured on two subscales of the Swanson, Nolan, and Pelham Rating Scale (SNAP- IV; Swanson, Nolan, & Pelham, 1992). The measure assesses symptoms of inattention and hyperactivity/impulsivity. We also included the SNAP-IV subscale oppositional defiant disorder (ODD). The first two subscales have 9 items; the third, 8; and responses are given on 4-point scales, indicating the frequency of the symptoms. The reliabilities for inattention and ODD were .91 at both pre- and posttest, and .92 and .91 for hyperactivity/impulsivity at pre- and posttest, respectively. To examine clinical relevance, we used prevailing 95th percentile cutoff values (see Swanson, n.d.).

Parent Outcome Measures

Parents' negative reactions. We assessed parents' negative reactions to child noncompliance and misbehavior, specifically *angry outbursts* and *harsh treatment*. The Angry Outbursts scale (Stattin, Persson, Burk, & Kerr, 2011) measures how parents respond when their child does something they "really do not like." The scale includes items like "My first reaction is anger and I yell at the child," and "I have problems controlling my irritation in such situations," Parents responded on a 3-point scale, ranging from 1 (*never*) to 3 (*most often*; α = .79 and .78 at pre- and posttest, respectively).

The Harsh Treatment subscale of the Parents Practice Interview (Webster-Stratton, Reid, & Hammond, 2001) was used to measure harsh parenting reactions in response to child's noncompliance. Parents indicated the frequency of their negative responses when the child "does something s/he is not supposed to do" on a 7-point scale, ranging from 1 (*never*) to 7 (*always*). The measure contains seven items, such as "Raise your voice (scold or yell)," and "Threaten to punish him/her (but not really punish him/her)" (α = .63 at pretest, and .72 at posttest).

Parents' positive reactions. We measured parents' *attempted understanding* (Stattin et al., 2011) of child's misbehavior. In relation to the stem question about how parents respond when their child does something they "really do not like," the Attempted Understanding subscale includes items such as "The most important thing to me is to understand why the child did what he or she did," and "I try to understand how the child thought and felt," all rated on a 3-point scale ranging from 1 (*never*) to 3 (*always*; α = .68 at both pre- and posttest).

We measured parents' use of rewards using a subscale of the Parents Practice Interview (Webster-Stratton et al., 2001). The subscale Rewarding the Child contains the items: "Buy something for him/her (such as special food, a small toy) or give him/her money for good behavior," and "Give him/her an extra privilege (such as cake, go to the movies, special activity for good behavior)," which were rated on a 7-point scale ($\alpha = .79$ at pretest, and .75 at posttest).

Parenting Competence and Well-Being

Parenting competence. We used the revised structure, by Gilmore and Cuskelly (2009), of the Parents' Sense of Competence measure (PSOC; Johnston & Mash, 1989) to assess parents' perceived satisfaction and efficacy. Six items measured parent's *satisfaction* with the parenting role, and five items parents' sense of *efficacy*. Parents responded to the items on a scale from 1 (*strongly disagree*) to 6 (*strongly agree*). The subscales were correlated at $r = .44$ at pretest, and $r = .52$ at posttest, and were combined into one broader parenting competence measure. The reliabilities of the overall scale were .81 at pretest, and .95 at posttest.

Stress. Parents' experienced *stress* was measured using the Caregiver Strain Questionnaire (Brannan, Heflinger, & Bickman, 1997). For the current study, we used the objective-strains subscale, which has 10 items. Parents were asked to report on how they had been affected by their children's problems during the past six months (e.g., How often were you interrupted in what you were doing? How often did you have to stay home from work or neglect other duties?). At posttest, we asked the parents to rate the same items with reference to the preceding month. The responses were rated on a scale ranging from 1 (*no time*) to 5 (*very often*; $\alpha = .90$ at both pre- and posttest).

Depressive symptoms. We used the 20-item Center of Epidemiological Studies—Depression Scale (Radloff, 1977) to assess the depressive symptoms of parents. The measure is widely used with normative populations, and has demonstrated similar psychometric properties to other well-validated measures, such as the Beck Depression Inventory (Shafer, 2006). Parents reported on whether they experienced each of the symptoms during the previous week on a 4-point scale, ranging from 1 (*not at all*) to 4 (*often*; $\alpha = .92$ at pretest, and .93 at posttest). We used the prevailing 95th percentile cutoff values that are applicable to normal samples of Swedish adults (Scott & Melin, 1998).

Demographic characteristics. Parents responded to a number of questions regarding the sociodemographic features of the primary reporter and her/his family. Age and gender were reported on in response to open-ended questions. *Marital status* was assessed by asking whether the primary informant was married, cohabiting, widowed, single, or something else. *Immigrant status was* assessed according to whether one or both parents had migrated to Sweden from a country outside Europe. Parents reported on where they had been born (in Sweden, another Scandinavian country, another European country, or a country outside Europe). Household *monthly income* was rated on a 6-point scale, from 1 (*0–10,000 SEK*), to 6 (*more than 50,000 SEK per month*). *Economic strain* was assessed according to whether income matched household expenditure. The reporter rated this on a 4-point scale, ranging from 1 (*Our monthly income does not cover our expenses*) to 4 (*Our monthly income is fine and we do not think about what we spend*). The primary reporter's highest educational level was measured on a 4-point scale ranging from 1 (*compulsory school*) to 4 (*university degree*).

Attendance and Implementation Fidelity

The numbers of sessions and meeting times differed between the programs: Comet, eleven 2.5-hr sessions; Cope and Incredible Years, twelve 2-hr sessions; Connect, ten 1-hr sessions. We calculated percentages for the parents attending program sessions. Overall, 70% of the parents attended at least 75% of all the sessions, 18.6% attended between 50% and 75% of the sessions, 4.9% attended between 20 and 50% of the sessions, and only 6.5% attended fewer than 25% of the sessions. Attendance was significantly higher for Comet (78.8%) and Connect (85.2%) than for Incredible Years (72.1%) and Cope (69.3%), $F(3, 442) = 13.34$, $p < .001$.

Parents rated their *satisfaction* with the programs on the basis of a single question: How do you like the parenting program that you have just finished?, using a 5-point scale ranging from 1 (*very bad*) to 5 (*very good*). Average satisfaction was very high ($M = 4.47$, $SD = .73$). Parents were significantly more satisfied with Comet (4.73) than with Incredible Years (4.51), Connect (4.36), and Cope (4.31), $F(3, 594) = 11.14$, $p < .001$.

We measured implementation fidelity through expert ratings of a randomly selected set of videotaped sessions. The sessions of each parenting group were videotaped on three randomly selected occasions, which generated 216 videotapes. We randomly selected 25% of these recordings ($n = 56$), stratified by the number of sessions for each program. Two specialists in each program, for example, people who had received training, who had implemented such a program, and who had trained group leaders, viewed the videotapes. They rated treatment fidelity (i.e., the extent to which the group leader followed the program manual) on a 10-point scale ranging from 1 (*not at all*) to 10 (*totally*). Interrater agreement, as indicated by the correlation between the ratings of the independent assessors was high ($r = .84$). The expert ratings of treatment fidelity were high, with an average rating for all programs of $M = 7.93$, and with mean ratings across the programs ranging from 6.86 to 9.04. Levene's F test, followed by a post hoc test, showed that the ratings of Cope (9.04) and Comet (7.94) were significantly higher than those of Connect (7.52) and Incredible Years (6.86). Overall, attendance of parents, parents' satisfaction, and implementation fidelity were all satisfactory.

Analyses

We used latent change models (Duncan, Duncan, & Strycker, 2013) to examine the changes in child and parent outcomes over time. It is possible to model measurement error in latent-variable models, which provide more robust estimates of change over time (Duncan et al., 2013). In a latent change model for two assessments, the intercept refers to the mean level of the measure at baseline, and the slope refers to the magnitude of change from baseline to posttest. To achieve model identification, we fixed the residuals to the measurement error rates computed on the basis of the interitem reliabilities of the measures at each time point (Duncan et al., 2013). To compare the effects of participating in each parenting program with the wait-list condition, we created four

dummy-coded variables, each representing one of the parenting programs, and used the wait-list condition as the reference category. The regression paths from each of the dummy-coded variables to the intercept factor enabled baseline differences across the programs to be tested. The paths from the dummy-coded variables to the slope factors provided a test of whether program participation led to greater change in children and parents compared with the wait-list condition. Next, we contrasted the programs (with each other) by comparing the magnitudes of the regression coefficients. These contrasts provided tests of whether the effectiveness of the programs (with reference to the wait-list condition) were significantly different from each other. Among the parenting programs, Connect is designed for older children, which resulted in differences in child age across the programs. Therefore, we controlled for the effect of child age in all models. Finally, we tested the moderating roles of age, child gender, family income, economic strain, immigrant origin, child use of concomitant medication, and parents' receipt of additional therapeutic services on the effectiveness of the parenting programs.

In the current data set, observations were clustered around four regions, 30 communities, and 72 parenting groups. In clustered data, variations in measurements may show substantial differences across clusters, resulting in biased estimates of parameters and inflated Type I error rates (Hox, 2010). Design effects for the outcome measures were computed using intraclass correlations (ICC) to establish whether clustering in the data needed to be taken into account in the analyses (Hox, 2010). The average ICC was .013 for regions, .06 for parenting groups, and .08 for communities. Clustering in an analysis is recommended when a design effect is larger than 2 (B. Muthén & Satorra, 1995). The design effects ranged from 1.0 to 1.3 for the regions and 1.1 to 1.4 for parenting groups, whereas the design effects across communities ranged between 1.02 and 2.40. Accordingly, clustering across communities was taken into account when estimating all the models using MPlus 7 software (L. K. Muthén & Muthén, 1998–2012).

To obtain unbiased estimates despite attrition, we employed intention-to-treat analysis. We adopted a conservative approach to handling of missing data by attributing baseline values to missing data at Time 2, an imputation procedure that assumes that participants with missing values at Time 2 did not change over time (Streiner & Geddes, 2001). Cohen's *d* effect sizes were computed to estimate the effects of the programs compared with the wait-list condition while taking the initial level of each outcome measure into account (Derzon, Sale, Springer, & Brounstein, 2005). We adopted the normal convention of interpreting a Cohen's *d* of .2 as a small effect, .5 as a medium effect, and .8 as a large effect.

Results

Differences in Participant Characteristics at Pretest

Before examining the program outcomes, we tested whether the randomization procedure had eliminated potential differences at baseline across conditions. As reported in Table 3, the children and parents who were allocated to the study conditions did not differ significantly, except on two demographic factors. The children and parents who took part in the Connect program were older than those on the other programs, which was to be expected because Connect was only delivered to parents of children age 8–12 years. Overall, 25% of the parents reported that they received additional therapeutic services, and 14% of the children were using medications related to their behavior problem at baseline. However, neither the distribution of children using medication, $\chi^2(3) = 7.021$, $p = .071$, nor the parents' receiving additional services, $\chi^2(3) = 5.93$, $p = .115$, were distributed significantly different across the program conditions. Regarding differences on the outcome measures at baseline, the only significant difference was that children on the Connect program had higher inattention/hyperactivity symptoms than those in the wait-list condition, $F(4, 783) = 3.55$, $p = .007$, $\eta^2 = .02$. Although the small effect size suggested that this difference was not substantial, we controlled for the effect of age in all the models. Overall, the randomization seemed to have ensured that parents and children on the four parenting programs and in the wait-list condition were relatively equivalent to each other when the programs began.

Changes in Children's Problem Behavior and ADHD Symptoms

Did parents' participation in the programs lead to changes in their children's behavior over time? Our primary child outcome

Table 3
Sociodemographic Characteristics of the Program Participants

Characteristic	Comet	Cope	Incredible Years	Connect	Wait-list control	$F/\chi^2(df)$	p
Child							
Child gender (*d*) (boy)	64.90%	61.10%	67.10%	66.30%	60.40%	2.34 (4)	.673
Child age	7.32_a (2.41)	7.07_a (2.54)	6.93_a (2.15)	9.8_b (1.35)	6.71_a (2.35)	59.25 (4, 774)	<.001
Parent							
Parent gender (*d*) (mother)	88.80%	86.90%	79.50%	81.60%	91.20%	10.85 (4)	.028
Parent age	37.96_a (6.63)	37.28_a (6.07)	36.9_a (5.95)	39.97_b (5.24)	37.41_a (5.98)	6.84 (4, 766)	<.001
Marital status (*d*) (married)	77.10%	71.80%	70.70%	71.10%	77.40%	3.44 (4)	.487
Immigrant status (*d*)	13.10%	14%	16%	12.60%	20.50%	5.22 (4)	.265
Monthly income	4.25 (1.54)	4.01 (1.65)	3.77 (1.60)	4.02 (1.64)	4.07 (1.44)	1.39 (4, 766)	.236
Economic strain	2.74 (0.64)	2.74 (0.71)	2.61 (0.78)	2.68 (0.81)	2.81 (0.66)	1.16 (4, 771)	.329
Education level	3.15 (0.95)	3.11 (1.07)	2.84 (0.95)	3.09 (1.02)	3.23 (0.91)	2.19 (4, 773)	.068

Note. Standard deviations of mean values are presented in parentheses. Same subscripts indicate nonsignificant between-groups differences and different subscripts indicate significant group differences.

measure came from the ECBI. Our secondary outcome measure was based on SNAP-IV, a measure of ADHD symptoms. For descriptive purposes, we first estimated the proportional reductions in mean values from pretest to posttest for all program and wait-list conditions. The average reduction in mean ECBI intensity and problems was 29% in the program condition (range = 21%–46%), and 16% (range = 11%–20%) in the wait-list condition. We observed that, for all child problem behaviors and on the ADHD symptom measure, there were 11% to 20% reductions among children in the wait-list condition. The reductions were not unexpected because parents were allowed to seek help from regular services. What was striking was the magnitude of the decrease among the program participants. The average reduction in mean ECBI intensity and problems was 29% in the program condition (range = 21%–46%), but only 16% (range = 11%–20%) in the wait-list condition. In parallel, the average reduction on the ADHD measures was 27% (range = 21%–35%) in the program group, but 21% (range = 15%–17%) in the wait-list group. Consistently, the latent change models showed that the problem behaviors and ADHD symptoms of the children whose parents participated in the programs decreased significantly more than those in the wait-list condition (see Table 4). The effect sizes of the reductions, as measured by Cohen's *d*, indicated medium effects of the programs with values ranging from .26 to .63 for ECBI intensity, and .17 to .35 for the ECBI problems measure. As for the SNAP-IV scales, the effect sizes of the reductions were smaller: d = .01–.17 for inattention, .10–.22 for hyperactivity, and .07–.26 for oppositional defiance disorder. Overall, participation in the parenting programs was related to a moderate decrease in child behavior problems, and a smaller but significant decrease in ADHD symptoms.

The comparisons of the magnitudes of the program effects suggested significant differences between the programs (see Table 4). Specifically, the greatest reduction in ECBI intensity scores was among the children of parents who participated in Comet (d = .63), which was followed by Cope (d = .44), the Incredible Years program (d = .42), and the Connect program (d = .31). Also, Comet (d = .49) was more effective than the other three programs in reducing ECBI problems scores (d = .27 for Cope and Incredible Years, and d = .17 for Connect). Overall, participants in Comet decreased most in child behavioral problems, followed by participants in Cope and in Incredible Years. Participants in Connect showed the lowest reduction in child behavioral problems.

The Connect program did not lead to any significant reduction in inattention, hyperactivity, or ODD symptoms. By contrast, Comet, Cope, and Incredible Years achieved similar reductions in both inattention and ODD symptoms. Regarding hyperactivity, Cope and Incredible Years resulted in significantly greater reductions than Comet and Connect, whose reductions did not differ from that of the wait-list condition. In sum, Comet, Cope, and Incredible Years achieved rather similar reductions in ADHD or ODD symptoms, but Connect was not effective in reducing these symptoms.

The Clinical Relevance of the Programs for Children

The clinical significance of the effects of prevention programs is of primary concern in research, practice, and policy. Making a strong argument for the effectiveness of parenting programs requires the demonstration of reductions in problematic behaviors and symptoms among individuals who are rated above clinically relevant levels at the start of a program. We examined whether the parenting programs achieved clinically significant reductions in child problem behaviors and ADHD symptoms. We identified the children in our sample who scored above the previously identified ECBI scale values of the 95th percentile in Swedish normative samples (Axberg et al., 2008). Similarly, we used norm values for the three SNAP-IV measures of ADHD and ODD symptoms to identify the children in our sample whose inattention, hyperactivity/impulsivity, and ODD symptoms were above the 95th percentile at baseline and postmeasurement (see http://www.ADHD.net/snapiv-instructions.pdf).

Overall, we observed that 59% to 78% of the children who scored higher than the 95th percentile in ECBI intensity at pretest were under this cutoff point at posttest. Similarly, 44% to 64% of those who scored over the 95th percentile for ECBI problems at pretest scored under this cutoff point at posttest. Results of the

Table 4
Changes in Child Outcomes and the Effect Size Estimates

Child outcome	Comet		Cope		Incredible Years		Connect	
	β	*d* [95% CI]	β	*d* [95% CI]	β	*d* [95% CI]	β	*d* [95% CI]
ECBI Intensity	$-.43_{a}^{***}$	.63 [.53, .71]	$-.30_{b}^{***}$	.44 [.36, .53]	$-.26_{b}^{***}$	.42 [.32, .52]	$-.21_{c}^{***}$	.31 [.21, .41]
≥95%	$-.14_{a}^{**}$	.32 [.28, .36]	$-.16_{a}^{***}$	.28 [.25, .32]	$-.21_{b}^{***}$	.35 [.30, .39]	$-.06_{c}$	.07 [.03, .11]
ECBI Problems	$-.24_{a}^{***}$	.49 [.46, .51]	$-.14_{b}^{**}$	.27 [.24, .29]	$-.16_{ab}^{***}$	.27 [.23, .29]	$-.11_{b}^{*}$	.17 [.14, .20]
≥95%	$-.19_{a}^{***}$	.49 [.44, .54]	$-.11_{b}^{*}$	.25 [.20, .30]	$-.16_{a}^{***}$	.26 [.20, .32]	$-.11_{b}^{*}$	.21 [.15, .26]
SNAP-IV Inattention	$-.15_{a}^{*}$	.17 [.10, .24]	$-.13_{a}^{*}$	.08 [.01, .15]	$-.18_{a}^{**}$	.18 [.09, .26]	$-.08_{b}$	.01 [.00, .08]
≥95%	−.07	.10 [.06, .14]	−.08	.00 [.01, .04]	−.07	.07 [.03, .11]	−.08	.03 [.00, .07]
SNAP-IV Hyperactivity	$-.09_{a}$	.15 [.07, .22]	$-.16_{b}^{**}$	.19 [.12, .27]	$-.15_{b}^{**}$	.22 [.14, .30]	$-.03_{a}$	.10 [.02, .18]
≥95%	−.06	.17 [.14, .23]	$-.10^{*}$	.19 [.15, .24]	$-.14^{**}$	.28 [.23, .34]	−.08	.19 [.16, .25]
SNAP-IV ODD	$-.19_{a}^{***}$	.26 [.19, .33]	$-.14_{a}^{**}$	.23 [.16, .30]	$-.15_{a}^{**}$	.25 [.17, .32]	$-.07_{b}$	.07 [.01, .14]
≥95%	−.05	.06 [.03, .09]	.10	.06 [.03, .08]	−.06	.12 [.09, .16]	−.04	.07 [.03, .10]

Note. Standardized beta coefficients represent the effect of each program (vs. the wait-list control condition) on the changes in the outcome measure from baseline to post-test measure. The Cohen's *d* estimates are the effect size of each program (vs. the wait-list control condition) at posttest on child outcome. Different subscripts of beta coefficients indicate significant differences in the effect of the respective program in the reductions of child outcomes. ECBI = Eyberg Child Behavior Inventory (Colvin, Eyberg, & Adams, 1999); SNAP-IV = Swanson, Nolan, and Pelham Rating Scale (Swanson, Nolan, & Pelham, 1992); ODD = oppositional defiant disorder; PSOC = Parents' Sense of Competence measure (Johnston & Mash, 1989).
* $p < .05$. ** $p < .01$. *** $p < .001$.

modeling showed that child participants in Comet, Cope, and Incredible Years showed significantly more clinically relevant reductions in ECBI intensity (*d*s = .32, .28, .35, respectively) and on the ECBI problems measure (*d*s = .49, .25, .26, respectively) than did the children in the wait-list condition (see Table 4). By contrast, participants in Connect did not show significantly more clinically relevant reductions on these measures than those in the wait-list condition. Next, we found 41% to 59% reductions in attention symptoms from pretest to posttest among the children who scored above the level of the 95th percentile at pretest. The reductions, again using the clinical cutoffs, were 35% to 46% for hyperactivity, and 33% to 59% for ODD symptoms. However, these reductions were not consistently greater than the reductions observed in the wait-list condition. Specifically, only participation in Cope (*d* = .19) and Incredible Years (*d* = .28) showed a higher clinically relevant reduction in hyperactivity symptoms. In sum, compared with the wait-list condition, participation in Comet, Cope, and Incredible Years, but not in Connect, led to clinically relevant reductions in child problem behaviors relative to being on the wait-list. As for ADHD symptoms, only the reductions in hyperactivity among children whose parents participated in Cope and Incredible Years were greater than among the children whose parents were on the wait-list.

Changes in Parents' Reactions, Parenting Competence, and Well-Being

We examined the effects of participating in the programs in relation to four aspects of parenting: negative reactions to children, positive parenting behaviors, parental competence, and parents' psychological well-being. Prior to making statistical comparisons, we inspected the percentage decrease in these parenting outcomes from pretest to posttest. Overall, the wait-list condition showed 0% to 12% changes in mean values, whereas the parents who participated in the programs showed changes in the range 4% to 30%, averaged across the four programs. The smallest change was a 4% increase in positive parenting, and the greatest a 30% decrease in depressive symptoms, among the parents who had scored above the level of the 95th percentile.

The latent change models suggested greater interprogram variations in changes on the parent outcome measures than on child outcomes, where the effects tended to be similar (see Table 5). Specifically, the parents who participated in Comet showed significantly higher reductions in negative parenting behaviors, such as angry outbursts (*d* = .30) and harsh parenting (*d* = .58), relative to those in the wait-list condition. Also, parents who participated in Cope decreased in harsh parenting (*d* = .39) more than did those on the wait-list. The changes in positive parenting practices, such as attempted understanding and use of rewards were less consistent. First, involvement in a program did not increase parents' attempted understanding when they witnessed child misbehavior. On the other hand, participants in Comet, Incredible Years, and Connect increased in their use of rewards as a parenting strategy (*d*s = .30, .30, and .28, respectively). One of the most consistent changes in parent outcomes was related to sense of parenting competence. Parents on all four programs significantly increased in perceived parenting competence. The effect sizes for the programs were medium, but there was significantly greater change in the participants in Comet (*d* = .69) than in the participants in Cope, Incredible Years, and Connect (*d*s = .47, .32, and .35, respectively).

Parents who participated in the programs also showed improvements in their psychological well-being, despite some differences across the programs. Specifically, there were significant decreases in stress among the parents who participated in Comet (*d* = .30), Incredible Years (*d* = .23), and Connect (*d* = .13). In addition, participants in Comet (*d* = .38), Cope (*d* = .36), and Connect (*d* = .30) decreased significantly in depressive symptoms, with similar effect sizes across the programs. We also examined decreases in clinical levels of depressive symptoms by testing whether parents who scored above the level of the 95th percentile on the depressive symptoms measure at baseline scored lower than the 95th percentile at posttest. The changes in the clinical levels were in line with

Table 5
Changes in Parent Outcomes and Effect Size Estimates

Parent outcome	Comet β	Comet *d* [95% CI]	Cope β	Cope *d* [95% CI]	Incredible Years β	Incredible Years *d* [95% CI]	Connect β	Connect *d* [95% CI]
Negative parenting								
Angry Outbursts	$-.29_{a}^{***}$	.30 [.25, .35]	$-.11_{b}$	.16 [.12, .22]	$-.11_{b}$	.12 [.06, .18]	$-.02_{b}$	.10 [.05, .15]
Harsh Treatment	$-.26_{a}^{**}$	.58 [.52, .53]	$-.18_{b}$	.39 [.33, .45]	$-.15_{b}$	.28 [.21, .35]	$-.09_{b}$	.18 [.11, .25]
Positive parenting								
Attempted understanding	.15	.14 [.10, .18]	.16	.23 [.19, .27]	.04	.09 [.05, .13]	.10	.27 [.24, .31]
Rewards	$.18_{a}^{**}$	.30 [.17, .44]	$.10_{b}$	.25 [.11, .39]	$.14_{a}^{**}$	.30 [.15, .46]	$.19_{a}^{**}$	.28 [.13, .43]
Parental competence								
PSOC	$.35_{a}^{***}$	.69 [.62, .76]	$.25_{b}^{***}$	.47 [.40, .54]	$.15_{b}^{**}$	.32 [.24, .40]	$.24_{b}^{***}$	.35 [.28, .42]
Mental health								
Stress	$-.17_{a}^{**}$	.30 [.24, .36]	$-.07_{b}$	.10 [.03, .17]	$-.16_{a}^{***}$	.23 [.15, .31]	$-.11_{ab}^{*}$	.13 [.05, .21]
Depression	$-.19_{a}^{***}$	.38 [.31, .44]	$-.20_{a}^{***}$	.36 [.29, .42]	$-.07_{a}$	.20 [.13, .28]	$-.15_{a}^{***}$	.21 [.15, .28]
Depression ≥ 95th	$-.14_{a}^{**}$	.30 [.25, .35]	$-.13_{a}^{*}$	.24 [.19, .28]	$-.04_{a}$	.10 [.05, .16]	$-.14_{a}^{*}$	.23 [.18, .28]

Note. Standardized beta coefficients represent the effect of each program (vs. the wait-list control condition) on the changes in the outcome measure from baseline to posttest measure. The Cohen's *d* estimates are the effect size of each program (vs. the wait-list control condition) at posttest on parent outcome. Different subscripts of beta coefficients indicate significant differences in the effect of the respective program in the reductions of parent outcomes. PSOC = Parents' Sense of Competence measure (Johnston & Mash, 1989).
$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

the changes in the mean levels of the symptoms. Specifically, there were significant decreases in the clinical levels of depressive symptoms among the parents who participated in Comet ($d = .30$), Cope ($d = .24$), and Connect ($d = .23$), and the magnitudes of the changes were not statistically different between the programs.

In sum, we found that the parent training programs achieved reductions in negative parenting behaviors and improvements in positive parenting practices, parenting competence, and psychological well-being. However, the effects varied across the programs, with Comet leading to the most consistent changes across all parent outcomes, except parents' attempts to understand child misbehaviors. The second most effective program, with five significant changes among the eight parent outcomes, was Connect. Cope and Incredible Years yielded significant changes on only three out of the eight parent outcome measures.

Intention-to-Treat Analysis

Attrition is inevitable in any study designed to include follow-up. Loss of participants to follow-up may lead to the drawing of biased conclusions regarding the effectiveness of preventive interventions, especially in real-world conditions (Gartlehner et al., 2006). Thus, we employed an intention-to-treat analysis.

Overall, the findings of the intention-to-treat analysis did not change our results concerning program effects, except with regard to estimates of three specific child outcomes. The effects of Connect on change in the mean level of the ECBI problems scale and change among children scoring above the 95th percentile cutoff from pretest to posttest became nonsignificant. As a result, on the basis of the intention-to-treat analysis, we can argue that the Connect program was only effective with regard to changes in ECBI intensity. Also, the significant effect of Cope on child inattention symptoms disappeared. It should be noted that all these effect estimates were significant at the $p < .05$ level in the initial analyses, albeit with small effect sizes. Regarding parent outcomes, all the previously significant effects estimates again reached significance in the intention-to-treat analysis. We also estimated effect sizes using the intention-to-treat data. This approach led to changes in the effect size estimates only to the second or third decimal, which indicates that the estimated effects of the programs on both child and parent outcomes were highly stable.

Moderators of Program Effects

To test whether the effects of the parenting programs varied across child and parent groups with different demographic characteristics, we ran a series of moderated regression models. We specifically tested whether child's age and gender, parent's age, family income, economic strain, marital status, immigrant origin, child use of medication, and parents' receipt of additional therapeutic services moderated program effects on the children and parents. The combination of a wide array of child and parent outcomes and seven moderating factors generated a large number of models. Specifically, we fitted 70 separate models to test the effects of moderating factors on child outcomes, including the models testing the clinical relevance of program effects. And, in addition, we fitted 56 models for parent outcomes. Among all of the interaction terms, 5% of unique interaction terms for child outcomes, and 4.5% for the parent outcomes were significant at $p < .05$. To account for the risk of an inflated Type I error rate due to a large number of unique interaction terms and models being tested sequentially, we set the significance level for the main and interaction effects of the moderators at $p = .0007$ for child outcomes, and $p = .0009$ for parent outcomes. It proved that none of the main effects of the seven moderators or their interaction terms reached significance. The lowest p value we observed was for the interaction effect between economic strain and the Connect program in predicting changes in parents' use of rewards ($p = .001$), whereas the second lowest p value was for the interaction between economic strain and the Comet program in predicting changes in children's hyperactivity ($p = .004$). We also estimated the effect sizes for all moderators, and unique interaction terms. Overall, the mean effect size estimate was $f^2 = .009$, and the median effect size estimate was $f^2 = .007$. Among the unique statistically significant interaction terms, both the mean and median effect size was $f^2 = .01$. In addition, none of these significant interaction terms were systematically accumulated across a given moderator or program. In sum, the low rate of significant estimates among a large number of tested interaction effects, low effect sizes of the significant estimates, and inconsistent patterns of these effects across the moderators and programs raise concern about the reliability of these findings. Therefore, we did not interpret these interaction effects because they were likely to arise by chance. Overall, the findings suggest that the parenting programs we evaluated had stable effects across children and parents with different sociodemographic characteristics.

The findings related to the use of concomitant medication deserve special attention because of its implication for practice. Only three of the interaction terms that involve use of medication were significant. Specifically, use of medication moderated the effect of Cope on ECBI Problem scale ($f^2 = .014$, $p = .020$); the effect of Connect on ECBI Problem scale ($f^2 = .009$, $p = .033$); and the effect of Comet on SNAP-IV Inattention subscale ($f^2 = .016$, $p = .015$). Given the small magnitude of the effect sizes, and the large number of interaction terms tested, these observed effects might not indicate a systematic role of use of concomitant medication on the effectiveness of programs.

Finally, in order to make overall comparisons between the four programs, we controlled for child's age, which was necessary because Comet, Incredible Years, and Cope were used for children ages 3 to 8 years, and Comet, Cope, and Connect for children ages 9 to 12 years. We also performed separate analyses for children who were between 3 and 8 years of age, and those between 9 and 12 years. These analyses, which treated the younger and the older children separately, essentially reproduced what we have presented here for the broader age group (of 3 to 12 years). Hence, the findings obtained for the children ages 3 to 12 were about the same as those for the 3- to 8-year-olds and the 9- to 12-year-olds separately.

Discussion

A primary goal of the present study was to evaluate the short-term effectiveness of three behavioral programs and one nonbehavioral program. Our aim was to understand the impacts of these evidence-based programs on parents who had profound difficulties in handling their children's behavior problems, and who turned to the human service units available in their communities for help. We conducted an effectiveness evaluation

and examined the changes that took place between pretest (1–2 weeks before the programs were started) and posttest (about 3–4 months after the programs began).

First, we found that all the four programs were more effective than the wait-list in reducing children's conduct problems and ADHD symptoms. Second, we found that the programs had adequate clinical relevance, in that they reduced behavior problems among the children with problem levels above the 95th percentile cutoff at pretest. When it comes to effects on parents, we found that parents who took part in a parenting program expressed a much greater sense of efficacy in their parental role, and, except for parents attending Incredible Years, decreased their levels of stress and depression more than those who received treatment as usual. Moreover, parents on the Comet, Incredible Years, and Connect programs increased over time in adopting rewarding behaviors, compared with parents receiving treatment as usual. We also found differences between the programs. Participants in Comet showed less negative behaviors, such as angry outbursts and harsh treatment, toward their children at posttest, compared with the parents who received treatment as usual. The other programs did not have such an effect. Moreover, few changes in children's ADHD symptoms were found for the parents attending the Connect program. Finally, effect sizes differed considerably between the programs, with Comet generally being the most effective program for changing children's behaviors. When it comes to the parent measures, Comet and Connect were the most effective, whereas Cope and Incredible Years were the least effective. Our study extends the results of earlier efficacy trials to an effectiveness trial, and suggests that parenting programs are more effective than the treatment that is usually offered in real-life settings.

A further goal of our study was to examine possible differences in effectiveness between the programs. It is interesting that both types of programs, the behavioral and the nonbehavioral, improved children's conditions and enhanced parenting, but there were some substantial differences. Comet was the most effective, and Connect was the least effective, with Cope and Incredible Years in the middle. Although some systematic reviews have reached the conclusion that behavioral treatments achieve somewhat better results than nonbehavioral treatments (Serketich & Dumas, 1996), Lundahl et al. (2005) did not find any statistically significant differences between them. These meta-analyses, however, have compared programs that are quite different in terms of study populations, measures used, and sociocultural contexts. In our study, by contrast, by comparing four commonly employed parenting programs in one and the same trial, we were able to show that both nonbehavioral and behavioral parenting programs can help parents and children in the short term, but that the behavior-based programs are somewhat to be preferred. An explanation for the differences between the types might be that nonbehavioral programs need more time to show their effects, because they are aimed at changing relationships dynamics, and modifying relationships might take longer than it takes directly to influence behaviors. This might result in better outcomes for behavioral parenting programs than nonbehavioral programs in the short term, but not necessarily in the long run.

An unexpected result was the difference found in effects between two of the behavioral programs, Comet and Incredible Years. Both programs are strongly influenced by Patterson's coercion model. They both seek to enhance parents' abilities to positively communicate with their children, primarily through rules and limit-setting, and through praising and rewarding the child's prosocial behaviors, and ignoring the child's problem behaviors. Indeed, Comet was explicitly inspired by Webster-Stratton's (1993) ideas about selective prevention when the program was developed. Despite these similarities with regard to theoretical foundations, Comet was generally more successful in reducing children's behavioral problems than Incredible Years. There are several possible explanations for this. First, Incredible Years had lower implementation fidelity than Comet, and fidelity in implementation has been associated with larger intervention effects (Durlak & DuPre, 2008). Thus, the differences between Comet and Incredible Years might be due to variable quality in implementation. Second, we are unaware of any cultural adaptation of Incredible Years when it was introduced into Sweden, but we do know that Comet was specifically developed for Swedish families. The developers put in extensive efforts to create a program that maintained the principal features of Patterson's Parent Management Training and Incredible Years while, at the same time, remaining culturally appropriate for Swedish families (Kling, Sundell, Melin, & Forster, 2006). For example, all the materials, examples, and visual presentations were thoroughly evaluated by parents and social service professionals working with parents and children, and then, the program was revised and improved based on their feedback (Kling et al., 2006). It is likely that these improvements made the program content more relevant for parents, and easy to embrace the suggested strategies for managing their child's behavior. Even though empirical studies have not yet been conducted, it has been theorized that the cultural adaptation of a parenting program is an important aspect of understanding its success (e.g., Castro, Barrera, & Martinez, 2004; Sundell, Ferrer-Wreder, & Fraser, 2013). Thus, the differences in effects between the two programs might be due to different levels of cultural adaptation. Moreover, there were some problems in organizing sufficiently large Incredible Years groups in two of the communities we considered, with the consequence that some parents needed to travel further than parents in the other programs to take part. Finally, the parents who attended the Comet program were significantly more satisfied than parents who took part in Incredible Years. In short, a number of reasons—concern with fidelity, cultural adaptation, organizational problems, and parental satisfaction—might explain why two very similar behavioral programs ultimately had different effects. We cannot say which of these factors was the most important in making Incredible Years less effective than Comet.

Our study demonstrated that the sociodemographic characteristics of parents and their children are unlikely to affect the effectiveness of parenting programs of the kind examined in this evaluation. However, meta-analyses, mostly conducted in North American settings, suggest that parenting programs are less effective among economically disadvantaged families, and among families with single and depressed mothers (Lundahl et al., 2006; Reyno & McGrath, 2006). The findings of our study are more in line with a recent British study that found no moderating effect of socioeconomic status in a replication of the Incredible Years program (Gardner, Hutchings, Bywater, & Whitaker, 2010). We might hypothesize that the macrosocietal context (in America, economic disparities in the population are larger than in Europe)

may play a role in determining the importance of participants' sociodemographic characteristics for program effectiveness.

Our findings also demonstrated that the impact of the parenting programs on children's behavioral outcomes were rather robust regardless of whether the children were using concomitant medication. This finding is contrary to the "common sense" expectation: use of concomitant medication to treat disruptive behaviors and ADHD symptoms should boost the program effects. The current study is limited in answering thoroughly the question of whether medication or parenting programs, or the combined use of medication and parenting programs, is more effective in helping children with conduct problems. However, the conclusion is clear. Use of medication neither increased nor decreased the effectiveness of the parent training programs. There has been an increasing trend, since 1990s, in use of medication to treat children and adolescents who display symptoms of conduct disorders and ADHD both in North America and Europe (e.g., Olfson, Blanco, Liu, Moreno, & Laje, 2006; Hsia & Maclennan, 2009; Steinhausen & Bisgaard, 2014). Our findings imply that effective evidence-based parenting programs may be a viable alternative to use of medication. Further research is needed to formally test relative and combined effectiveness of medication and parent training programs to reach stronger conclusions for practice and public health policy.

In the analyses where we tested for program effects, the programs were tested against the wait-list condition, representing treatment as usual. Treatment as usual in the current study was found, from pre- to posttest, to lead to fewer behavioral problems among children, better parenting of the children, and better parental health. The changes in the wait-list group were significant at the .001 level, for ECBI intensity and problems, and for the SNAP-IV symptom scales measuring inattention, hyperactivity, and oppositional defiance. In short, there were substantial changes in the wait-list group from pretest to posttest. How do we explain the fact that changes took place in the wait-list as well as in the program condition?

Important findings of the study are that treatment as usual is not synonymous with not receiving any treatment at all, and that taking part in a parenting program does not mean that parents do not get additional help elsewhere. The parents who were randomized to the various program and wait-list conditions in this evaluation appear to have received substantial help from the communities in which they lived. This might explain the substantial changes also observed for individuals on the wait-list. We know that one in four of all the parents received extra help in school or at home from the community at pretest. At posttest, the figure was one in five. This continuous community help was delivered equally frequently to the parents who took part in a parenting program as to the parents who were on the wait-list. In short, many of our parents received help from the community, irrespective of whether they had been randomized to a parenting program or to the wait-list. Some of these parents were invited to take part in parent groups, where they were given advice, and were encouraged to reflect and practice over three to four months. It is this "extra" effects generated by the parenting programs that we examine in this study. Because program effectiveness was not influenced by the use of medication related to behavior problems, one implication for professionals could be to consider evidence-based parent training a viable first-hand alternative for families with child behavior problems. Inattentive, restless, or oppositional behaviors among some of the children, which are usually the reason for considering medication, might be reduced with parent training programs that teaches behavior management strategies, such as use of praise and incentives, ignoring children's inappropriate behaviors, and positive discipline. Medication was not included as a specific treatment condition in the present study design. Thus, we need more knowledge on whether and to which extent psychotropic medication enhances the effects of parent training for subgroups of conduct problems, and children with certain risk factors or behaviors. These aspects should be addressed in future randomized controlled trials.

Strengths and Limitations

The major limitation of this study lies in its reliance on self-reporting. Data were collected through parental self-reports, with no external or independent sources for evaluating child or parent behaviors. Because of the multicenter design, with several agencies involved, it was not possible to arrange for observations of parent–child behaviors. Thus, future studies should use a multi-informant design in attempts to confirm our results.

The study has important strengths. To start with, the present study is the very first trial to evaluate several parenting programs, both behavioral and nonbehavioral, with the same study design and setting. It is a large, well-powered, randomized multicenter controlled trial, with very low attrition. Because of the powerful design, we can be fairly confident of the results. We also managed to enroll the parents of children with various levels of severity of behavior problems, which allowed us to conduct analyses of those above the 95th cutoff for conduct problems or ADHD. By so doing, we were able to test the clinical relevance of the programs. A further strength of the study is that we collected information about adherence to treatment manuals, and evaluated treatment fidelity as well as session quality through the ratings of videos by independent experts on each program.

Implications for Research and Practice

Meta-analyses and reviews have established that parenting programs are generally effective in the short term in reducing child disruptive behaviors (e.g., Furlong et al., 2012; Serketich & Dumas, 1996). Also, based on various levels of scientific evidence, clinical and systematic evaluations have identified well-established and probably efficacious programs, which are then recommended to clinicians (Eyberg, Nelson, & Boggs, 2008). The present study adds to more specific knowledge about the effects of some international and national behavioral and nonbehavioral programs by evaluating them in the same context in an effectiveness trial. We conclude that parenting programs provide help beyond what parents and children can already obtain in the normal health care system, independent of the sociodemographic characteristics of the parents and children. Overall, the behavioral programs were found to be more effective than the nonbehavioral program in reducing children's problem behaviors (comparing Comet, Incredible Years, and Cope, with Connect). The results of the present study might suggest that there are mechanisms common to different parenting programs. However, it might also be the case that different mechanisms are important according to the natures of the behavioral programs, and that observations of different mechanisms are particularly important when comparing behavioral and nonbehavioral programs. The next step is to illuminate more specifically the mecha-

nisms that contribute to effects, so that programs can be operationalized according to evidence-based mechanisms of change. This might make therapists more aware of the changes that make a difference for a specific program, and what they, as therapists, need to achieve to improve the chances of achieving successful outcomes if these changes (or criteria) are not observed.

References

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: Author.

Axberg, U., & Broberg, A. G. (2012). Evaluation of "The Incredible Years" in Sweden: The transferability of an American parent-training program to Sweden. *Scandinavian Journal of Psychology, 53,* 224–232. http://dx.doi.org/10.1111/j.1467-9450.2012.00955.x

Axberg, U., Johansson Hanse, J., & Broberg, A. G. (2008). Parents' description of conduct problems in their children—A test of the Eyberg Child Behavior Inventory (ECBI) in a Swedish sample aged 3–10. *Scandinavian Journal of Psychology, 49,* 497–505. http://dx.doi.org/10.1111/j.1467-9450.2008.00670.x

Bandura, A. (1986). *Social foundations of thought and Action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Barkley, R. A. (1997). *Defiant children: A clinician's manual for assessment and parent training*. New York: Guilford Press.

Brannan, A., Heflinger, C. A., & Bickman, L. (1997). The Caregiver Strain Questionnaire measuring the impact on the family of living with a child with serious emotional disturbance. *Journal of Emotional and Behavioral Disorders, 5,* 212–222. http://dx.doi.org/10.1177/106342669700500404

Castro, F. G., Barrera, M., Jr., & Martinez, C. R., Jr. (2004). The cultural adaptation of prevention interventions: Resolving tensions between fidelity and fit. *Prevention Science, 5,* 41–45. http://dx.doi.org/10.1023/B:PREV.0000013980.12412.cd

Cedar, B., & Levant, R. F. (1990). A meta-analysis of the effects of parent effectiveness training. *American Journal of Family Therapy, 18,* 373–384. http://dx.doi.org/10.1080/01926189008250986

Colvin, A., Eyberg, S., & Adams, C. (1999). *Restandardization of the Eyberg Child Behaviour Inventory*. Unpublished manuscript.

Costin, J., & Chambers, S. M. (2007). Parent management training as a treatment for children with oppositional defiant disorder referred to a mental health clinic. *Clinical Child Psychology and Psychiatry, 12,* 511–524. http://dx.doi.org/10.1177/1359104507080979

Cunningham, C. (2005). Large group, community based, family-centered parent training. In R. A. Barkley & K. R. Murphy (Eds.), *Attention deficit hyperactivity disorder: A clinical workbook* (pp. 480–498). New York, NY: Guilford Press.

Cunningham, C. E., Bremner, R., & Boyle, M. (1995). Large group community-based parenting programs for families of preschoolers at risk for disruptive behaviour disorders: Utilization, cost effectiveness, and outcome. *Journal of Child Psychology and Psychiatry, 36,* 1141–1159. http://dx.doi.org/10.1111/j.1469-7610.1995.tb01362.x

Derzon, J. H., Sale, E., Springer, J. F., & Brounstein, P. (2005). Estimating intervention effectiveness: Synthetic projection of field evaluation results. *The Journal of Primary Prevention, 26,* 321–343. http://dx.doi.org/10.1007/s10935-005-5391-5

Duncan, T. E., Duncan, S. C., & Strycker, L. A. (2013). *An introduction to latent variable growth curve modeling: Concepts, issues, and application*. Mahwah, NJ: Erlbaum.

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41,* 327–350. http://dx.doi.org/10.1007/s10464-008-9165-0

Eyberg, S. M., Nelson, M. M., & Boggs, S. R. (2008). Evidence-based psychosocial treatments for children and adolescents with disruptive behavior. *Journal of Clinical Child and Adolescent Psychology, 37,* 215–237. http://dx.doi.org/10.1080/15374410701820117

Farmer, E. M. Z., Compton, S. N., Bums, J. B., & Robertson, E. (2002). Review of the evidence base for treatment of childhood psychopathology: Externalizing disorders. *Journal of Consulting and Clinical Psychology, 70,* 1267–1302. http://dx.doi.org/10.1037/0022-006X.70.6.1267

Furlong, M., McGilloway, S., Bywater, T., Hutchings, J., Smith, S. M., & Donnelly, M. (2012). Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years. *Cochrane Database of Systematic Reviews, 2,* CD008225.

Gardner, F., Hutchings, J., Bywater, T., & Whitaker, C. (2010). Who benefits and how does it work? Moderators and mediators of outcome in an effectiveness trial of a parenting intervention. *Journal of Clinical Child and Adolescent Psychology, 39,* 568–580. http://dx.doi.org/10.1080/15374416.2010.486315

Gartlehner, G., Hansen, R. A., Nissman, D., Lohr, K. N., & Carey, T. S. (2006). *Criteria for distinguishing effectiveness from efficacy trials in systematic reviews*. Rockville, MD: Agency for Healthcare Research and Quality. Retrieved from http://www.ncbi.nlm.nih.gov/books/NBK44029/

Gilmore, L., & Cuskelly, M. (2009). Factor structure of the Parenting Sense of Competence scale using a normative sample. *Child: Care, Health and Development, 35,* 48–55. http://dx.doi.org/10.1111/j.1365-2214.2008.00867.x

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications*. New York, NY: Routledge.

Hsia, Y., & Maclennan, K. (2009). Rise in psychotropic drug prescribing in children and adolescents during 1992–2001: A population-based study in the United Kingdom. *European Journal of Epidemiology, 24,* 211–216. http://dx.doi.org/10.1007/s10654-009-9321-3

Johnston, C., & Mash, E. J. (1989). A measure of parenting satisfaction and efficacy. *Journal of Clinical Child Psychology, 18,* 167–175. http://dx.doi.org/10.1207/s15374424jccp1802_8

Kling, A., Forster, M., Sundell, K., & Melin, L. (2010). A randomized controlled effectiveness trial of parent management training with varying degrees of therapist support. *Behavior Therapy, 41,* 530–542. http://dx.doi.org/10.1016/j.beth.2010.02.004

Kling, Å., Sundell, K., Melin, L., & Forster, M. (2006). Komet för föräldrar. *En randomiserad effektutvärdering av ett föräldraprogram för barns beteendeproblem. FoU-rapport, 14*. [Comet for parents. A randomized efficacy evaluation of a parental program for children with behavioral problems. Research&Development report, 14]

Lindsay, G., & Strand, S. (2013). Evaluation of the national roll-out of parenting programmes across England: The parenting early intervention programme (PEIP). *BMC Public Health, 13,* 972. http://dx.doi.org/10.1186/1471-2458-13-972

Litschge, C. M., Vaughn, M. G., & McCrea, C. (2010). The empirical status of treatments for children and youth with conduct problems: An overview of meta-analytic studies. *Research on Social Work Practice, 20,* 21–35. http://dx.doi.org/10.1177/1049731508331247

Little, M., Berry, V., Morpeth, L., Blower, S., Axford, N., Taylor, R., & Tobin, K. (2012). The impact of three evidence-based programmes delivered in public systems in Birmingham, United Kingdom. *International Journal of Conflict and Violence, 6,* 260–272.

Lundahl, B., Risser, H. J., & Lovejoy, M. C. (2005). A meta-analysis of parent training: Moderators and follow-up effects. *Clinical Psychology Review, 26,* 86–104. http://dx.doi.org/10.1016/j.cpr.2005.07.004

Maughan, D. R., Christiansen, E., Jenson, W. R., Olympia, D., & Clark, E. (2005). Behavioral parent training as a treatment for externalizing behaviors and disruptive behavior disorders: A meta-analysis. *School Psychology Review, 34,* 267–286.

Menting, A. T., Orobio de Castro, B., & Matthys, W. (2013). Effectiveness of the Incredible Years parent training to modify disruptive and proso-

cial child behavior: A meta-analytic review. *Clinical Psychology Review, 33,* 901–913. http://dx.doi.org/10.1016/j.cpr.2013.07.006

Michelson, D., Davenport, C., Dretzke, J., Barlow, J., & Day, C. (2013). Do evidence-based interventions work when tested in the "real world?" A systematic review and meta-analysis of parent management training for the treatment of child disruptive behavior. *Clinical Child and Family Psychology Review, 16,* 18–34. http://dx.doi.org/10.1007/s10567-013-0128-0

Moretti, M., Holland, R., Moore, K., & McKay, S. (2004). An attachment-based parenting program for caregivers of severely conduct disordered adolescents: Preliminary findings. *Journal of Child and Youth Care Work, 19,* 170–179.

Moretti, M. M., & Obsuth, I. (2009). Effectiveness of an attachment-focused manualized intervention for parents of teens at risk for aggressive behaviour: The Connect Program. *Journal of Adolescence, 32,* 1347–1357. http://dx.doi.org/10.1016/j.adolescence.2009.07.013

Muthén, B., & Satorra, A. (1995). Complex sample data in structural equation modelling. *Sociological Methodology, 25,* 267–316. http://dx.doi.org/10.2307/271070

Muthén, L. K., & Muthén, B. O. (1998–2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Author.

National Institute for Clinical Excellence. (2006). *Conduct disorder in children—parent-training/education programmes*. Retrieved from http://www.nice.org.uk/TA102

Nixon, R. D. V. (2002). Treatment of behavior problems in preschoolers: A review of parent training programs. *Clinical Psychology Review, 22,* 525–546. http://dx.doi.org/10.1016/S0272-7358(01)00119-2

Obsuth, I., Moretti, M. M., Holland, R., Braber, K., & Cross, S. (2006). Conduct disorder: New directions in promoting effective parenting and strengthening parent-adolescent relationships. *Canadian Academy of Child and Adolescent Psychiatry, 15,* 6–15.

Olfson, M., Blanco, C., Liu, L., Moreno, C., & Laje, G. (2006). National trends in the outpatient treatment of children and adolescents with antipsychotic drugs. *Archives of General Psychiatry, 63,* 679–685. http://dx.doi.org/10.1001/archpsyc.63.6.679

Patterson, G. R. (1982). *Coercive family process*. Eugene, OR: Castalia Publishing Co.

Patterson, G. R., Dishion, T. J., & Chamberlain, P. (1993). Outcomes and methodological issues relating to treatment of antisocial children. In T. R. Giles (Eds.), *Handbook of effective psychoterapy*. New York, Plenum Press.

Piaget, J., & Inhelder, B. (1962). *The psychology of the child*. New York: Basic Books.

Radloff, L. S. (1977). The CES-D scale a self-report depression scale for research in the general population. *Applied Psychological Measurement, 1,* 385–401. http://dx.doi.org/10.1177/014662167700100306

Reid, M. J., Webster-Stratton, C., & Hammond, M. (2003). Follow-up of children who received the Incredible Years intervention for oppositional-defiant disorder: Maintenance and prediction of 2-year outcome. *Behavior Therapy, 34,* 471–491. http://dx.doi.org/10.1016/S0005-7894(03)80031-X

Reyno, S. M., & McGrath, P. J. (2006). Predictors of parent training efficacy for child externalizing behavior problems—A meta-analytic review. *Journal of Child Psychology and Psychiatry, 47,* 99–111. http://dx.doi.org/10.1111/j.1469-7610.2005.01544.x

Scott, B., & Melin, L. (1998). Psychometric properties and standardised data for questionnaires measuring negative affect, dispositional style and daily hassles. A nation-wide sample. *Scandinavian Journal of Psychology, 39,* 301–307. http://dx.doi.org/10.1111/1467-9450.00088

Serketich, W. J., & Dumas, J. E. (1996). The effectiveness of behavioral parent training to modify antisocial behavior in children: A meta-analysis. *Behavior Therapy, 27,* 171–186. http://dx.doi.org/10.1016/S0005-7894(96)80013-X

Shafer, A. B. (2006). Meta-analysis of the factor structures of four depression questionnaires: Beck, CES-D, Hamilton, and Zung. *Journal of Clinical Psychology, 62,* 123–146. http://dx.doi.org/10.1002/jclp.20213

Stattin, H., Persson, S., Burk, W. J., & Kerr, M. (2011). Adolescents' perceptions of the democratic functioning in their families. *European Psychologist, 16,* 32–42. http://dx.doi.org/10.1027/1016-9040/a000039

Steinhausen, H. C., & Bisgaard, C. (2014). Nationwide time trends in dispensed prescriptions of psychotropic medication for children and adolescents in Denmark. *Acta Psychiatrica Scandinavica, 129,* 221–231. http://dx.doi.org/10.1111/acps.12155

Streiner, D., & Geddes, J. (2001). Intention to treat analysis in clinical trials when there are missing data. *Evidence-Based Mental Health, 4,* 70–71. http://dx.doi.org/10.1136/ebmh.4.3.70

Sundell, K., Ferrer-Wreder, L., & Fraser Mark, W. (2014). Going global: A model for evaluating empirically supported family-based interventions in new contexts. *Evaluation & The Health Professions, 37,* 203–230. http://dx.doi.org/10.1177/0163278712469813

Swanson, J., Nolan, W., & Pelham, W. E. (1992). The SNAP-IV rating scale. Retrieved from http://www.adhd.net

Swanson, J. (n.d.). *Sample SNAP-IV Teacher and Parent Rating Scale #6160*. Retrieved from http://www.myadhd.com/snap-iv-6160-18sampl.html

Thorell, L. B. (2009). The Community Parent Education Program (COPE): Treatment effects in a clinical and a community-based sample. *Clinical Child Psychology and Psychiatry, 14,* 373–387. http://dx.doi.org/10.1177/1359104509104047

van den Hoofdakker, B. J., van der Veen-Mulders, L., Sytema, S., Emmelkamp, P. M. G., Minderaa, R. B., & Nauta, M. H. (2007). Effectiveness of behavioral parent training for children with ADHD in routine clinical practice: A randomized controlled study. *Journal of the American Academy of Child & Adolescent Psychiatry, 46,* 1263–1271. http://dx.doi.org/10.1097/chi.0b013e3181354bc2

Webster-Stratton, C. H. (1993). Strategies for helping early school-aged children with oppositional defiant and conduct disorders: The importance of home-school partnerships. *School Psychology Review, 22,* 437–457.

Webster-Stratton, C., Hollinsworth, T., & Kolpacoff, M. (1989). The long-term effectiveness and clinical significance of three cost-effective training programs for families with conduct-problem children. *Journal of Consulting and Clinical Psychology, 57,* 550–553. http://dx.doi.org/10.1037/0022-006X.57.4.550

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2001). Preventing conduct problems, promoting social competence: A parent and teacher training partnership in head start. *Journal of Clinical Child Psychology, 30,* 283–302. http://dx.doi.org/10.1207/S15374424JCCP3003_2

Webster-Stratton, C., Reid, M. J., & Hammond, M. (2004). Treating children with early-onset conduct problems: Intervention outcomes for parent, child, and teacher training. *Journal of Clinical Child and Adolescent Psychology, 33,* 105–124. http://dx.doi.org/10.1207/S15374424JCCP3301_11

Received February 20, 2014
Revision received March 16, 2015
Accepted March 23, 2015 ■