

**EARLY CHILDHOOD CONSULTATION PARTNERSHIP:
RESULTS OF A RANDOM-CONTROLLED EVALUATION**

FINAL REPORT AND EXECUTIVE SUMMARY

Walter S. Gilliam

May 15, 2007

Contact Information

Walter S. Gilliam, PhD
Assistant Professor of Child Psychiatry and Psychology
Child Study Center
Yale School of Medicine
230 South Frontage Road
PO Box 207900
New Haven, CT 06520-7900

Phone: 203-785-3384

Email: walter.gilliam@yale.edu

Acknowledgements

This evaluation was supported by a grant from the Child Health and Development Institute of Connecticut. Judith Meyers provided considerable intellectual support for this evaluation and its dissemination. Crista Marchesseault coordinated the data collection and assisted with describing the ECCP service model. Liz Bicio and Sam Moy, both of Advanced Behavioral Health, were invaluable partners for the information they openly shared about the ECCP model and its implementation, as well as their willingness to discuss the emerging findings and potential program implications. Appreciation is also due to the many ECCP consultants and teachers who agreed to be a part of this evaluation.

Early Childhood Consultation Partnership: Results of a Random-Controlled Evaluation

EXECUTIVE SUMMARY

This report summarizes the findings of a random-controlled crossover evaluation of the Early Childhood Consultation Partnership (ECCP), a statewide system of early childhood mental health consultation in Connecticut. A detailed version is provided in the Final Report.

BRIEF DESCRIPTION OF ECCP

Created in 2002 by a combination of public and private funds, ECCP is an early childhood mental health consultation system available to staff at child care centers serving young children (infants to 5-years old) throughout Connecticut. Typically, services are requested by child care center directors or staff, where there are either behavioral or social-emotional concerns for individual children or classroom-wide behavioral management challenges. The ECCP service model is brief but relatively intensive – 8 weeks long, with 4 to 6 hours of classroom-based consultation per week provided by one of ten supervised masters-level consultant supported by ECCP, plus a week-12 follow-up visit. The intervention is loosely manualized and menu-driven based on individualized needs of teachers and classrooms. In addition to providing teacher training on various behavioral and social-emotional topics, the consultation has two main areas of focus (described in detail in the final report): (a) Classroom-specific consultation – focusing on improving teacher-child interactions, classroom behavior management, and overall program quality, and (b) child-specific consultation – focusing on improving teacher classroom behavioral and social-emotional strategies, parent partnerships, and community service referrals for specific children.

THE CURRENT EVALUATION

ECCP was evaluated in a randomized crossover evaluation design. During Cohort 1 (January 2005-June 2005), ECCP referrals were randomized to treatment ($n = 23$) or waitlist-control ($n = 23$) conditions. During Cohort 2 (September 2005-March 2006), 20 of the 23 waitlist-control classes accepted services and formed the treatment condition ($n = 20$), and classes were randomly selected from new referrals to form a no-treatment comparison group ($n = 19$). Across both cohorts, classrooms were mostly in community-based child care centers (82%), with the rest in Head Start centers (13%) or public schools (5%). Teachers completed child behavior ratings at pretest and posttest for the two children in each class whose behaviors were most concerning to the teacher during the pretest interval ($n = 144$, after attrition). Evaluation hypotheses were that participation in ECCP services would be associated with the following four outcomes:

1. Reduced teacher-rated behavior problems in target children (i.e., the two children in each classroom whose challenging behaviors concerned the teacher most),
2. Improved classroom environments and teacher-child interactions,
3. Increased teacher beliefs and practices regarding developmentally appropriate and child-centered pedagogy, and
4. Decreased teacher job stress and depression and increased teacher sense of job control and satisfaction.

EFFECTS ON TEACHER-RATED CHILD BEHAVIOR PROBLEMS

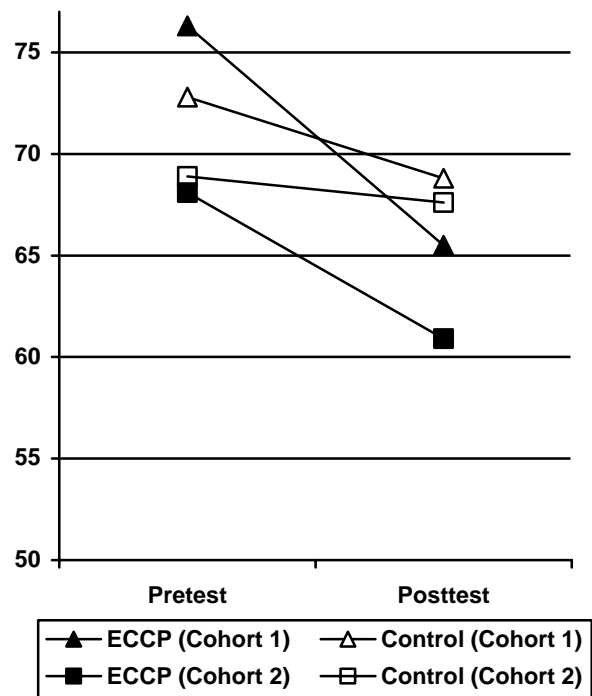
Controlling for baseline differences at the pretest interval, significant effects for ECCP were found on all scales of externalizing (acting-out) behavior problems on the two measures used in this evaluation. However, no effects were found for internalizing behavior problems (e.g., shyness, withdrawn behaviors, anxiety, etc.), social relatedness, or positive social skills.

In the Cohort 1 design where a randomized experiment was employed, children receiving ECCP showed decreases in teacher-rated oppositional behaviors that were significantly and meaningfully greater than what was found for the control group ($F_{(1,71)}=4.73$, $p<.05$, $d=0.53$). Effects were greater for hyperactivity ($F_{(1,71)}=6.83$, $p<.05$, $d=0.64$). The findings were replicated in Cohort 2 for oppositional behaviors ($F_{(1,67)}=5.77$, $p<.05$, $d=0.59$), but not for hyperactivity.

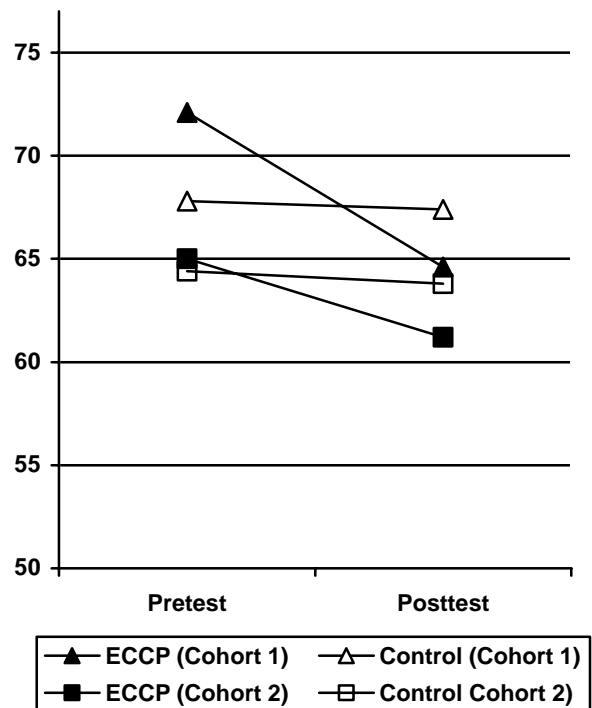
EFFECTS ON CLASSROOM QUALITY AND TEACHER-CHILD INTERACTIONS

Classroom environmental quality and teacher-child interactions were measured with the *Early Childhood Environment Rating Scale-Revised* (ECERS-R) and the *Caregiver Interaction Scale* (CIS) at posttest. All ratings were made by trained, statistically reliable, objective raters who were blind as to whether the classrooms they were rating had or had not received ECCP services. No treatment-control posttest differences were found for any of seven domains of classroom quality measured by the ECERS-R for either cohort. Similarly, no differences were found on the three teacher-child interaction scales from the CIS measuring amount of positive interactions, degree of teacher detachedness, or degree of teacher punitiveness for either cohort.

ECCP Effects on Oppositional Behaviors



ECCP Effects on Hyperactivity



EFFECTS ON TEACHER BELIEFS AND FEELINGS

Controlling for baseline differences at the pretest interval, no significant differences were found for either cohort on the four scales measuring teachers' (a) beliefs about the importance of obedience, (b) beliefs about the importance of child independence, (c) beliefs about developmentally appropriate classroom practices, or (d) self-reported developmentally appropriate classroom practices they use in their classrooms. Also, no significant differences for either cohort were found for teacher job stress, job control, job satisfaction or self-reported depression.

IMPLICATIONS

Overall, ECCP demonstrated statistically significant decreases in teacher-rated externalizing or acting-out behavior problems in the classroom, relative to the control group. Effects were greatest and of moderate magnitude in the areas of oppositional behaviors and hyperactivity. Findings did not support significant effects in reducing internalizing behavior problems (e.g., anxiousness, shyness, emotional lability) or improving social skills. Also, no significant effects of ECCP were found for observable classroom quality, teacher-child interactions, teacher beliefs regarding discipline or classroom management, or teacher job stress and satisfaction.

Given positive effects on child behavior, the lack of significant effects for classroom quality, teacher-child interactions, and teacher beliefs and feelings is puzzling. As a consultation service, ECCP is an indirect model of service. Consultative services were provided to the teachers, but not directly to children. No evidence, however, was found to support any of the hypothesized pathways of effect (through improving classroom environments, changing teacher beliefs, or reducing teacher job stress and depression). Therefore, exactly *how* ECCP is effective at reducing teacher-rated externalizing behavior problems remains unknown.

Whether the positive effects for reducing externalizing behavior problems would be observable to a trained, objective, rater who was blind to whether the classroom was receiving ECCP services was not measured in this evaluation and remains unknown. It is possible that the positive effects on children's behavior is based largely on teacher perceptions that may not be detectable to raters who are unaware of whether the classroom had received ECCP services or not. Treatment condition blind raters were used for the classroom quality and teacher-child interaction measures, where no treatment effects were found. Further evaluations should use a variety of informants for child behavior impacts, including teacher report, parent/caregiver report, and classroom behavioral observations.

Although much has been written about the provision of mental health consultation and related services in early education settings, there is little published evidence of the effectiveness of early childhood mental health consultation, as opposed to specific manualized techniques. As such, this evaluation is the first random-controlled evaluation of a widely implemented system of early childhood mental health consultation. Although the pathways of effect remain unclear in this indirect model of consultation, ECCP was effective at improving teacher-rated externalizing behavior problems in target children.

Early Childhood Consultation Partnership: Results of a Random-Controlled Evaluation

FINAL REPORT

Severe behavior problems during the preschool years are a meaningful predictor of continued behavior problems, poor peer standing, and academic difficulties during kindergarten (Howes, Calkins, Anastopoulos, Keane, & Shelton, 2003; Keane & Calkins, 2004). Fortunately, high quality early education and intervention programs may prevent severe behavior problems in young children from low-income communities and families (H. Yoshikawa, 1995; Zigler, Taussig, & Black, 1992). Yet some preschoolers may begin their early education programs with severe behavioral problems already present, potentially limiting their ability to participate fully and benefit from the early educational experience (Boyd et al., 2005).

Early childhood mental health consultation (ECMHC) may be an effective means for reducing severe behavior problems in early education and child care settings, as well as decreasing the likelihood of children with challenging classroom behaviors from losing services through expulsion and suspensions (Gilliam & Shahar, 2006). Recent findings from a national study of prekindergarten teachers indicated that teachers who report having an ongoing relationship with a classroom-based mental health consultant are about half as likely to report expelling a preschooler, relative to teachers who report no such support (Gilliam, 2005). Only 23% of these teachers, however, reported regular classroom access to a mental health consultant.

The lack of clear descriptions of ECMHC and compelling evidence of effectiveness has been a severe limitation for the field. Although the practice of ECMHC has been described in detail (Donahue, Falk, & Provet, 2000; Perry, Kaufmann, & Knitzer, 2007), there currently are no published reports of the effectiveness of ECMHC using rigorous evaluation methods. In a comprehensive review of all studies of ECMHC conducted between 1985 and 2005, 31 published and unpublished investigations of ECMHC were identified (Brennan, Bradley, Allen, Perry, & Tsega, 2005). Only 11 studies used quasi-experimental methods, and none involved random-controlled experiments. Overall, modest improvements in teacher- and parent-reported child social skills and behavior problems were reported, as well as improvements in teacher self-efficacy and parental involvement. Findings in the area of reduced job stress and improved job satisfaction were inconsistent, perhaps due to the use of various measures of unknown validity. However, differences in ECMHC model, intensity, and duration, as well as cases where the ECMHC was embedded as a part of a larger array of services, severely limit generalization, and none of the studies employed rigorous experimental designs capable of documenting ECMHC effects.

Only two of the studies reviewed by Brennan and colleagues (one quantitative and one qualitative) were ever reported in the peer-reviewed literature. In the only peer-reviewed *quantitative* study of ECMHC (Alkon, Ramler, & MacLennan, 2003), longer duration and higher intensity of consultation services were associated with lower staff turnover, higher teacher-reported self-competence, and higher program quality. However, sample sizes were small (9 centers receiving the service for more than one year compared to 14 centers receiving the service for under one year). In the only peer-reviewed *qualitative* study of ECMHC (Green, Simpson,

Everhart, Vale, & Gettman, 2004), staff who reported a good relationship with their mental health consultant was more likely to report that the consultant was better integrated into the program, had more clearly delineated roles, and was in general more effective. The study employed interview techniques in three Head Start grantees in the Pacific Northwest. Neither of these two studies was able to yield conclusions about the overall effectiveness of ECMHC, and neither included measures of child outcomes.

The purpose of this study is to evaluate rigorously the effectiveness of a statewide system of ECMHC on classroom quality and teacher-child interactions, teacher instructional beliefs and practices, teacher job stress and satisfaction, and child classroom behavior. A brief description of the ECMHC program being evaluated is provided below.

DESCRIPTION OF ECCP

Created in 2002 by a combination of public and private funds, the Early Childhood Consultation Partnership (ECCP) is an ECMHC program available to staff at child care centers serving young children (infants to 5-years old) throughout Connecticut. Typically, services are requested by child care center directors or staff, where there are either behavioral or social-emotional concerns for individual children or classroom-wide behavioral management challenges. The consultation focuses on the overall social-emotional atmosphere within the classroom, also addressing both behavioral concerns for individual children and classroom-wide behavioral management challenges.

The ECCP service model is 8 weeks long, with 4 to 6 hours of classroom-based consultation per week provided by one of ten supervised masters-level consultant supported by ECCP, plus a week-12 follow-up visit. The intervention is loosely manualized and menu-driven based on individualized needs of teachers and classrooms. In addition to providing teacher training on various behavioral and social-emotional topics, the consultation has two main areas of focus:

- **classroom-specific consultation**, focusing on improving teacher-child interactions, classroom behavior management, and overall program quality, and
- **child-specific consultation**, focusing on improving teacher classroom behavioral and social-emotional strategies, parent partnerships, and community service referrals for specific children.

During ECCP's first nine months of operation (October 2002 through June 2003), the program served 128 children and 72 classrooms at a total operating cost of \$605,886. During the second year, the number of children and classrooms served and operating costs increased. Child services and operating costs were relatively stable during years two through four, while the number of classrooms served decreased slightly. In Year 2 (July 2003 through June 2004), 198 children and 111 classrooms were served at an operating cost of \$975,542. In Year 3 (July 2004 through June 2005), 181 children and 91 classrooms were served at a cost of \$905,474. In Year 4 (July 2005 through June 2006), 206 children and 89 classrooms were served at a cost of

\$943,010. The current evaluation was conducted during ECCP's third and fourth years of operation.

ECCP Goals and Program Development

The goal of ECCP is to serve as an early prevention and intervention program for children birth to five; providing support, networking, and training opportunities in the areas of social and emotional health in early childhood, at family, community, and statewide levels. ECCP's service model was based on extant ECMHC literature (Donahue et al., 2000; Knitzer, 2000, 2001; Powell, Fixsen, & Dunlap, 2003) and principals found in various published curricula and measures (e.g., *Creative Curriculum for Preschool*, *Devereux Early Childhood Assessment Program*, and *Starting Small: Teaching Tolerance in Preschool and the Early Grades*). Programmatic materials (including an overview of the resources used in the development of ECCP; consultant orientation, training, and supervision plans; referral and service guidelines; details on programmatic components and assessment measures; and instructions for using the centralized data collection system) are housed at the centralized program management office that administers the program.¹ The manual is provided to each ECCP consultant and revised annually to reflect any program improvement changes.

ECCP Referral Process

The site referral process for ECCP is typically informal and comes from the ECCP consultants' active membership in various early childhood collaborative committees, advisory board memberships, community agencies, and other community-based groups, such as regional child care director meetings. Referrals for ECCP services are accepted from early education and child care centers serving children birth to five-years old, as well as parents or guardians of children exhibiting challenging behaviors. On average, about a quarter to a fifth of all ECCP referrals are from sites serving infants and toddlers, with the remainder mostly serving preschoolers (ages 3- to 5-years).

Typically, a site director or teacher calls the ECCP consultant assigned to their area directly to request services after hearing about the program in the community or through one of the local collaboratives with which ECCP participates. Referral sources can include parents, center directors or other administrators, and teachers. About two thirds of the referrals for the evaluation come from directors or other administrators. (There were no parent referrals for the study sites during the current evaluation.) ECCP consultants are assigned based on the geographical area of the referred centers.

ECCP Service Components

Classroom-specific consultation. Once a site referral is accepted, the consultation process begins with a meeting including the consultant assigned to the referred site, the site director or other administrator, and the classroom teacher(s). This meeting is followed by a classroom observation, upon which the consultant bases a classroom assessment such as the

¹ For more information, contact ECCP Program Manager Elizabeth Bicio, LCSW, at Advanced Behavioral Health (Phone: 860-704-6198; Email: ebicio@abhct.com). Ms. Bicio is responsible for the coordination and supervision of consultation services, the provision of training modules and initiatives, collaborating with mental health providers and community agencies; and other activities surrounding the overall oversight and management of the program.

Early Childhood Environmental Rating Scale—Revised (ECERS-R), the Infant/Toddler Environment Rating Scale (ITERS), and the Caregiver Interaction Scale (Arnett, 1989). The consultant then meets with the teacher again to gather additional information related to the teacher's strengths, specific areas of concern, and goals for the consultation. The ECCP consultant, then, enters all collected data into the centralized data collection system and produces a computer generated *Core Classroom Action Plan*, including a summary of classroom/teacher strengths, goals and objectives. The consultant then conducts a classroom team meeting to review and revise the plan with the teacher and program director or other administrator.

Based on the program model, the consultant spends 4-6 hours per week with each classroom/teacher during the eight weeks of regular consultation, providing technical assistance and assisting in the overall implementation of the action plan. Typically, 1 hour per week is spent in a team meeting with the teaching staff in the classroom during quiet moments (e.g., nap time) and 3 to 5 hours per week in the classroom working with the teaching staff. The exact amount of time spent in the classroom varies by teacher need, classroom structure, available resources, and consultant availability. Directors or other administrators are encouraged to participate in this process, though this level of participation varies and is at the discretion of the individual administrators.

On-going monthly mental health consultation groups held regionally by each ECCP consultant are also made available to all staff at participating sites, and a one-time training session is provided for each participating site by that site's ECCP consultant, who has been trained through a statewide professional development system to deliver specific social-emotional development trainings. These training sessions consist of a 90-minute workshop on a topic chosen by each site. More than 30 topic choices are offered, and ECCP consultants assist teachers in their topic selection. Example topics included *Verbal Environment: Talking to Children in Supportive Ways*; *Behavior Modification and Positive Reinforcement*; and *Proactive Approaches to Behavior Management*.

Child-specific consultation. Typically, about two-thirds of teachers in ECCP referred classrooms identify one to two children for child-specific consultation, with about a third of the teachers identifying no specific children. Target children are identified either prior to referral or with the ECCP consultant in the initial team meeting process. Parents/caregivers are contacted to obtain informed written consent for services prior to child-specific consultation. Although parents/caregivers may decline child-specific consultation, behavioral problems can still be addressed globally in the context of the classroom-specific consultation described above.

Once parent/caregiver consent is obtained, child-specific information is collected through demographic and referral information forms, informal conversations with the child's parents/caregivers and teacher, informal observations of the child in the classroom, and through the use of a variety of standardized teacher-rated measures of child social-emotional functioning and behavior. Computer-generated *Child Action Plans* are then created for each target child, including a section on observed strengths, as well as areas of concern and consultative goals identified through collaboration among the consultant, teachers, and parents/caregivers. The consultant then conducts a meeting with parents and teachers to review and revise the *Child*

Action Plan, to discuss classroom strategies for dealing with identified behavioral challenges, and to provide community-based resources and referrals for the child and family as needed.

ECCP Consultants

ECCP services are delivered by a total of ten consultants. Although the service was managed by a centralized program management organization, each of these consultants was employed individually by various subcontracted community-based agencies across Connecticut. At the time of the current evaluation's start, all ten consultants held a Master's degree in a human services related field, mainly in psychology or social work. All were also trained mental health clinicians, and although they were not required to be licensed mental health providers, six of the ten consultants held or were eligible for clinical professional licenses in counseling, marriage and family therapy, or clinical social work. About half had experience working in clinical mental health settings, about a third had worked in some capacity as mental health coordinators for Head Start, and the remainder had some amount of experience in the field of early childhood education. Nearly all of the consultants had worked with children under the age of six years, and about two-thirds had worked in education or child care settings prior to becoming an ECCP consultant. Five consultants had 2-3 years experience as an ECCP consultant, three had one year's experience, and two were new to ECCP.

Consultant recruitment. The centralized management organization overseeing ECCP recruited the consultants through their employing agencies (e.g., child guidance and mental health clinics). One main advertisement was distributed to subcontracting agencies in various communities around the state. Each of the consultants was employed by each of ten subcontracted community agencies. Each agency had its own geographic service area within Connecticut, as an aggregate covering the entire state.

Consultant training. Consultants received a variety of training lessons, organized into thirteen training units (the field of child care, family day care and kith-and-kin care, assessing quality care, child mental health, health promotion, children with special needs, abuse and neglect, adult learning, adult resiliency, consultation, team building, partnering with systems, and community planning). All consultants received specific training on the administration of all ECCP measures. Additional specific training topics included early childhood mental health consultation, multidisciplinary consultation, abuse/neglect petition filing, computer training, maintaining confidentiality and federal regulations, helping young children cope with trauma, managing aggression in the classroom, personal safety, cultural competency, psychotropic medication use in children, managing aggression in the classroom, and attachment. Community-based content experts provided these specific trainings.

Consultant supervision. Consultants were provided regular clinical supervision through three means: (a) ECCP group supervision, (b) ECCP individual supervision, and (c) agency-based supervision. The ECCP program manager provided the ECCP group and individual clinical supervision, whereas clinical supervisors at the consultants' employing community-based agency provided the agency-based supervision. The ECCP program manager was a Licensed Clinical Social Worker (LCSW), with a master's degree in social work and a bachelor's degree in special education specializing in social-emotional disorders. At the time of the evaluation, the program manager had a total of 15 years of experience in mental health

services, 10 of which were spent working in some capacity with young children from birth to five years old in both clinic- and school-based settings. The program manager was supervised by a licensed psychologist with over 20 years of clinical and management experience. No information is available on the qualifications of the agency-based supervisors.

ECCP group supervision was provided bi-weekly with each session being 2½ to 3 hours in length. These meetings included all ten ECCP consultants, were located at a centralized location, and focused on clinician-initiated case reviews for all ECCP consultations and provided ample opportunity for peer-to-peer learning, group supervision regarding the consultation work, and refreshers on all ECCP trainings previously discussed. *ECCP individual supervision* meetings were conducted with each consultant on a monthly basis, or more frequently as needed, and typically lasted for 2 hours each session. The purpose of the individual supervision was to provide an opportunity for individualized clinical consultation and review of clinical cases and data management, as well as to monitor individual consultant productivity and generate ideas with regard to the needs and direction of the ECCP program on the whole. *Agency-based supervision* was also provided to each ECCP consultant by clinical supervisors at each consultant's host agency. The frequency of this supervision varied by agency, ranging from weekly to monthly.

PREVIOUS ECCP PROCESS AND SATISFACTION EVALUATION

A process evaluation of ECCP implementation and global teacher ECCP satisfaction was conducted during ECCP's first year (Fink, Wakai, & Bruder, 2003). Although program maturity can increase the likelihood of documenting positive effects, process evaluations early in a program's development can be effective for improving program fidelity and locating problem areas to be addressed as the program matures to a point where an outcome evaluation is warranted (Gilliam & Leiter, 2003; Gilliam, Ripple, Zigler, & Leiter, 2000).

Fink et al.'s process evaluation results indicated reasonably good fidelity to the ECCP project goals in terms of range of classroom services delivered and completed and the characteristics of the targeted service population (i.e., children with aggressive behaviors that have been targeted currently or in the past for individualized services, such as special education). However, the number of children provided child-specific intervention was far less than anticipated (45% of the anticipated amount of child services across the classrooms enrolled in intensive, core and child-specific treatments).

A teacher survey conducted specifically for the Fink et al. evaluation was used to assess teacher global satisfaction with the program and feelings about the potential impact of ECCP. Many teachers reported feeling improvements in the classroom and children. Most teachers (57%) reported "great improvement" in the quality of their classroom environments, activities and interactions. Noticeably smaller proportions reported "great improvement" in staff resilience (41%) and partnerships with families (43%). For child behavior, 81% reported "modest" to "great" improvement in the target children, and 78% reported this level of improvement in the class on the whole. Also, 76% reported improvement in their ability to identify children in need of mental health referral, and 88% reported feeling that ECCP reduces the likelihood of suspensions or expulsions. Most teachers reported feeling sustained benefits from ECCP one to five months after services.

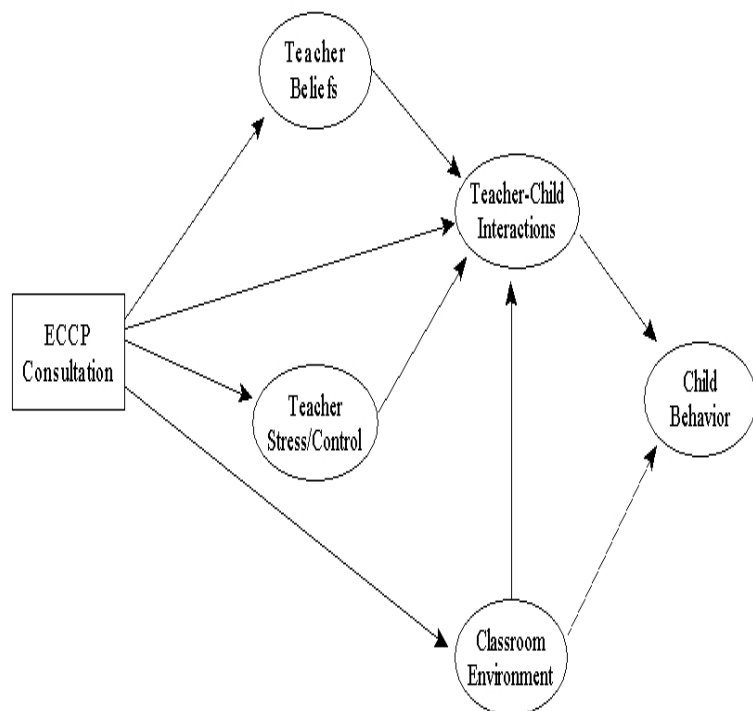
Although the overall findings for this first-year evaluation of ECCP implementation were positive, several limitations reduce their utility as an indicator of true program effectiveness. First, the response rate for a descriptive study was relatively low (44%), and the resulting small sample size ($n = 39$) reduces the faith one reasonably can place in these descriptive findings. Second, with neither (a) a no-treatment control or comparison group nor (b) a pretest of conditions before ECCP, it is impossible to know whether similar findings would have been noted even without ECCP. Third, the use of homemade measures of child behavior with global rating options (e.g., “modest improvement,” “great improvement,” etc.) and unknown reliability and validity are notoriously difficult to interpret. Nonetheless, this process and satisfaction evaluation was helpful in documenting general fidelity to service goals and encouraging hope in ECCP effectiveness.

THE CURRENT EVALUATION

The purpose of the current study was to evaluate the effectiveness of ECCP in a statewide methodologically rigorous random-controlled evaluation. ECCP implementation data were also collected, in order to describe the implementation context in which any effects might be found. As described above, ECCP is an indirect model of consultation, where the ECCP consultant primarily is working directly with the teacher, rather than the child or family. Therefore, any effects of ECCP on child behavior were hypothesized as resulting from changes in (a) the teacher’s knowledge and/or feelings about behavior management, (b) the quality of teacher-child interactions, (c) the quality of the classroom environment on the whole, and/or (d) the teacher’s feelings of stress and control when dealing with challenging behaviors, which might then impact teacher-child interactions or influence the way in which teachers’ view challenging behaviors.

Considerable research has demonstrated the effects of teacher-child interactions and classroom quality on children’s behaviors, and prekindergarten teachers who report elevated job stress and/or depression report nearly twice the rate of expulsions compared to teachers who do not report elevated job stress or depression (Gilliam, 2005; Gilliam & Shahar, 2006). Also, a small but statistically significant relationship between preschool teacher-reported depression and less sensitive teacher-child interactions has been reported in the literature (Hamre & Pianta, 2004). The hypothesized pathways of effect formed the basic logic model for the evaluation (as illustrated in the figure), and outcome measures were

ECCP EVALUATION LOGIC MODEL



selected with these factors in mind.

Evaluation hypotheses were that participation in ECCP services would be associated with the following four outcomes:

1. Increased teacher beliefs and practices regarding developmentally appropriate and child-centered pedagogy,
2. Decreased teacher job stress and depression and increased teacher sense of job control and satisfaction,
3. Improved classroom environments and teacher-child interactions, and
4. Reduced teacher-rated behavior problems in target children (i.e., the two children in each classroom whose challenging behaviors concerned the teacher most).

METHODS

ECCP was evaluated in a randomized crossover treatment design with two cohorts, employing both pretests and posttests. All pretest measures were completed within two weeks of the start of ECCP services (or what would have been the start for control classrooms), and all posttest measures were completed within two weeks of the final ECCP consultant visit (or what would have been the final visit for control classrooms). Evaluation Cohort 1 was recruited during ECCP's third year of operation (2004-2005), and Cohort 2 was recruited during ECCP's fourth year (2005-2006). Study inclusion criteria were that (a) the program must have a regularly meeting classroom-based component, and (b) the classroom must serve children predominately in the three- to five-year old age range, although children younger or older may also be served. The only exclusion criterion was that the classroom must not be located at a building-level site that had received ECCP services in the past. In order to ensure that all classrooms wanting ECCP services received them, the study was conducted when a modest ECCP waitlist had developed – treatment classrooms received the services immediately, while control classrooms were placed on a waitlist and received ECCP services within about three to four months.

During Cohort 1, 48 classrooms satisfied inclusion/exclusion criteria and were randomized to either immediately receive ECCP treatment ($n = 25$) or be placed on a waitlist to receive ECCP treatment during the following year ($n = 23$). Two treatment classrooms dropped out of the evaluation, yielding 23 treatment and 23 control classrooms during Cohort 1. During Cohort 2, the 23 waitlist-control classrooms from Cohort 1 were offered ECCP treatment, of which 20 agreed to receive ECCP treatment and participate in the evaluation. Additionally, 52 newly recruited classrooms during Cohort 2 satisfied inclusion/exclusion criteria, and 27 were randomly assigned to the control group, of which 19 agreed to participate in the study. This yielded 20 treatment and 19 control sites in Cohort 2. (See Table 1.)

SAMPLE

Across both cohorts, classrooms were mostly in community-based child care centers (82%), with the rest in Head Start centers (13%) or public schools (5%). Teacher and classroom characteristics for the treatment and control classes in both cohorts are presented in Table 2.

At the pretest interval for both cohorts, teachers were instructed to identify the two children (hereafter referred to as the *target children*) in their classrooms whose behaviors concerned them most, even if the teacher identified no such children to the ECCP consultant and even if ECCP services were not being directed toward these children. Only target children for whom both pretest and posttest data were collected were included in analyses (144 children of the potential 170). Of the 26 children for whom complete data were not obtained, valid pretest data were not obtained for 12, and posttest data were not obtained for an additional 14. Pretest data were not included for 10 children (Cohort 1 = 9, Cohort 2 = 1) for whom teacher ratings were not received during the pretest interval time frame, and pretest data for 2 children (both from one classroom in Cohort 1) were removed from analyses because the ECCP consultant mistakenly received the behavior rating forms from the teacher. Posttest data were not obtained for the 13 children (Cohort 1 = 6, Cohort 2 = 7) who withdrew from their early childhood program during the evaluation period, and 1 child (Cohort 1) for whom teaching ratings were not received within the posttest interval. Reasons for withdrawing from the program included withdrawal by the family ($n = 5$; family moved, convenience to parent workplace, unknown reasons), financial reasons ($n = 2$, failure to pay tuition, changes in employment or school status resulting in loss of tuition aid), changes in early education or child care provider ($n = 4$, transition to special education, home-based services, or a program for younger children), expulsion to home without services ($n = 1$), and unknown ($n = 1$).

See Table 3 for demographic characteristics of the 144 children for whom both pretest and posttest data were collected. There were no statistically significant differences between children for whom complete data were obtained versus those for whom complete data were not obtained in terms of either gender ($X^2_{(1)} = 1.19$, ns) or age at pretest ($F_{(1,156)} = 0.36$, ns).

ECCP SERVICE DURATION AND INTENSITY

ECCP treatment classes received a mean of 10.91 weeks of service ($SD = 1.56$, minimum = 8, maximum = 13). Within those service weeks, ECCP treatment classes received a mean of 43.07 service hours with a consultant ($SD = 10.02$, minimum = 24, maximum = 64). (See Table 4.) Amount of service hours received by ECCP classrooms varied considerably. Although the median number of total service hours was 41, 25% of classrooms received 34.75 hours or less and 25% of classrooms received 52.00 hours or more. Most classrooms received more than the ECCP model of 32 to 48 hours of services; however, 9.3% of the classrooms received less than 32 hours. The ECCP model was 4 to 6 service hours per week, and the mean number of hours received was 4.00 ($SD = 0.97$), ranging from a low of 2.42 hours to a high of 6.50 hours.

MEASURES

Evaluation measures were collected at both pre-treatment and post-treatment, assessing the overall quality of the child care environment, teacher beliefs and feelings, child behavior problems, and a variety of measures pertaining to the implementation of ECCP. All measures used in the evaluation are described below.

Classroom Measures

Measures of classroom quality were the *Early Childhood Environments Rating Scale-Revised* (Harms, Clifford, & Cryer, 1998) and the *Arnett Caregiver Interaction Scale* (Arnett,

1989). All classroom ratings are conducted by trained, reliable raters who were blind to treatment status. Raters received 1½ days of training on the ECERS-R and CIS, plus at least one practice administration in a classroom before rating in the field. Throughout data collection, each rater was paired with a different rater at a rate of at least one per six ratings to measure inter-rater reliability. Rater reliability measured in 16 classrooms was high for the ECERS-R total score ($ICC_{1,1} = .80$) and ranged from moderate (.47) to high (.89) across the seven ECERS-R subscales. Classroom measures were collected and reported at posttest only.

Early Childhood Environment Rating Scale—Revised (ECERS-R). The ECERS-R (Harms et al., 1998) is a well-known, standardized measure of the overall quality of child care environments. It is the most widely used instrument of its kind, and has been utilized extensively for both program evaluation and improvement. The ECERS-R has sound psychometric properties of reliability and validity as a measure of classroom quality with implications for child outcomes. The ECERS-R consists of 470 individual indicators of quality that contribute to 43 specific items located in 7 quality domains. The domains include (1) Space and Furnishings, (2) Personal Care Routines, (3) Language-Reasoning, (4) Activities, (5) Interaction, (6) Program Structure, and (7) Parents and Staff. Each item is scored on a seven-point anchored scale (1 = Inadequate; 3 = Minimal; 5 = Good; 7 = Excellent). Items within domains are averaged to yield a similarly interpreted score for each domain and for the total ECERS-R. Complete administration of the ECERS-R requires a well-trained rater and about 4 hours of classroom observation, followed by about 45 minutes of teacher interview.

Caregiver Interaction Scale (CIS). The CIS (Arnett, 1989) is another well-known instrument that measures quality of child care. Unlike the ECERS-R, which measures a wide variety of issues associated with child care quality, the CIS focuses exclusively on the quality of the interactions between the teacher(s) and children. The CIS consists of 26 items that measure four aspects of interaction: the frequency and quality of *positive interactions* and the degrees to which teachers are *not overly detached* from the children, *not overly permissive*, or *not overly punitive*. (The *not overly permissive* factor was omitted from the current evaluation, as this factor has been shown to be weak in validity and is commonly omitted in studies.) Each item is scored on a four-point scale and averaged to yield domain scores. Domain averages were transformed so that a high score always corresponded to something positive (e.g., high degrees of positive interactions, teachers not being too punitive, etc.).

Other measure of classroom environment. Classrooms were also rated with a measure of the overall classroom mental health climate. However, as this measure is still in development and has not yet been validated, no results are presented in this report.

Teacher Beliefs and Feelings Measures

Parental Modernity Scale (PMS). Teachers' beliefs about raising and educating young children were measured with the *Parental Modernity Scale of Child-Rearing and Education Beliefs* (PMS). The PMS assesses traditional authoritarian and progressive child-centered beliefs of parents and caregivers, with 30 items scored on a Likert scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*); higher scores indicate greater endorsement of each belief. Lower traditional scores and higher progressive scores have been shown to correlate positively with other positive caregiving behaviors and with children's intelligence and academic achievement

and negatively with child dependency and distractibility. Measures of internal (ranging from .88 to .94) and test-retest (.84) reliability are good (Schaefer & Edgerton, 1985). The PMS has been used in several large-scale studies of child care and state-funded prekindergarten to assess childrearing beliefs of child care caregivers and early childhood educators (Pianta et al., 2005).

Pre-K Survey of Beliefs and Practices (BP). The *Pre-K Survey of Beliefs and Practices* (Marcon, 1999) consists of 14 items (7 items measuring the child-centeredness of teacher beliefs and 7 items measuring the child-centeredness of teacher-reported classroom practices) that assess child-initiated versus academically-directed models of instruction. Higher scores (a greater level of child-initiated beliefs and practices) have been associated with greater child learning (Marcon, 1999). Teacher beliefs and practices are measured separately and scored on a continuum from 1 to 10.

Child Care Worker Job Stress Inventory (JSI). Teacher self-reported sense of job stress, control and satisfaction was measured with the 56-item *Child Care Worker Job Stress Inventory* (Curbow, Spratt, Ungaretti, McDonnell, & Breckler, 2000). The JSI consists of the three subscales that the authors hypothesized to be related to overall child care worker job stress: Job Control (how much control the teacher feels she or he has over things that occur at or around work), Job Demands (stressful situations and demands associated with providing early care and education), and Job Resources (things associated with providing early care and education that may help contribute to a teacher's job satisfaction or positive feelings about her or his own work; hereafter referred to as *Job Satisfaction*). All items are scored on a 5-point Likert scale, ranging from 1 (*rarely/never or very little*) to 5 (*most of the time or very much*). The JSI was validated on a sample of 196 randomly selected child care workers in Maryland. Cronbach alphas for the three original subscales ranged from .77 to .89. Concurrent validity studies, using other measures of job stress not specific to child care workers, demonstrated acceptable validity of this survey as a measure of child care related job stress.

Center for Epidemiological Studies Depression Scale (CES-D). Teachers reported the frequency of various symptoms of depression with the *Center for Epidemiological Studies Depression Scale* (Radloff, 1977). Respondents rated the frequency of 20 items on a four-point Likert scale, ranging from 0 (*rarely or none of the time; less than once per week*) to 3 (*most or all of the time; 5-7 days per week*). Sample items include "I was bothered by things that usually don't bother me" and "I felt that everything I did was an effort." The total CES-D score represents the sum of the 20 items. Scores may range from 0 to 60. Higher scores indicate more frequent depressive symptoms. Internal consistency for the total score ranges from .85 to .90, and test-retest coefficients average .57. The CES-D, one of the most widely used measures of depressive symptoms in adults, has indicated high known-groups and concurrent validity, and support has been found for its use across groups of different ethnicities.

Child Behavior Measures

Measures of child behavior problems for the two identified children in each class were the Conners' Teacher Rating Scale – Revised Long Form (Conners, 1997) and the Social Skills Rating System (Gresham & Elliot, 1990). Both of these measures have been used extensively in studies of behavior problems in preschoolers, and both are reliable and valid measures as rated

by preschool teachers and child care workers. Both the CTRS and SSRS were completed by the target children's lead classroom teacher at pretest and posttest intervals.

The CTRS yields five factor scores, plus two additional composite scores and a total score. The five factor scores include two externalizing behavior scores associated with (a) Oppositional Behaviors and (b) Hyperactivity, two internalizing behavior scores associated with (c) Anxious-Shy Behaviors and (d) Perfectionism, and a fifth factor associated with Social Problems. Two composite scales (Restless-Impulsive Behaviors and Emotional Lability) and a Total Behavior Problems score are also provided.

The SSRS measures both behavior problems as well as social skills. In terms of behavior problems, scores are provided for (a) Externalizing Behaviors, (b) Internalizing Behaviors, and (c) Total Behavior Problems. In terms of social skills, scores are provided for (a) Cooperation, (b) Assertion, (c) Self-Control, and (d) Total Social Skills.

In addition to the CTRS and SSRS, a measure of preschool expulsion risk was also used. However, this measure is still under development and is not yet validated. Therefore, results are not presented here.

PROCEDURES

Programs and classes were recruited through the normal ECCP channels and through fliers describing the evaluation study. The evaluators randomized classes to treatment and control conditions. For treatment classes, all pretest measures were collected within two weeks of the beginning of the intervention, and all posttest measures were collected within two weeks of the final ECCP consultation session. For control classes, pretest and posttest sessions were on a similar schedule equating the amount of time between pretest and posttest to that of the treatment group. Table 5 indicates which measures were administered at pretest and posttest intervals versus those administered only at posttest. Child and teacher measures were administered at pretest and posttest to account for naturally occurring changes, whereas classroom measures were considered to be more stable and were administered only at posttest.

ANALYSES

Data were analyzed separately for both Cohort 1 and Cohort 2. The Cohort 1 design is a random-controlled experiment, with classrooms randomly assigned by the evaluator to treatment and control conditions during Cohort 1 data collection. Cohort 2, however, is quasi-random assignment, because Cohort 2 treatment classes were not randomly assigned to treatment during Cohort 2 data collection but rather were assigned based on their Cohort 1 assignment to the waitlist control group. Data are also presented and analyzed with both cohorts combined. For classroom measures that were only administered during the posttest interval, simple ANOVA is used. For all other measures (where data were collected at both pretest and posttest intervals), ANCOVA is used to control for baseline (pretest) difference.

Because of the paucity of rigorously obtained findings for ECMHC in the literature from which specific outcome hypotheses can be generated, this evaluation is admittedly exploratory in nature. A broad array of potential outcome variables is included, and statistical analyses are conducted for each potential outcome separately.

Standardized measures of effect (Cohen's *d*) are provided for all analyses. For posttest-only analyses, Cohen's *d* was computed by subtracting the treatment group posttest mean by the control group posttest mean and dividing the difference by the pooled standard deviation. For analyses that included pretest scores, Cohen's *d* was computed by subtracting the covariance adjusted treatment group posttest mean by the covariance adjusted control group posttest mean and dividing the difference by the pooled covariance adjusted standard deviation (Cortina & Nouri, 2000). The direction of effect size is reversed for scales where a decrease in score is associated with a desired result (e.g., teacher job stress, child behavior problems, etc.). By convention, Cohen's *d* effect sizes were interpreted as trivial (below 0.20), small (0.20 to 0.49), moderate (0.50 to 0.79), or large (0.80 and above) (Cohen, 1988).

RESULTS

Results are presented below, organized by major outcome group: (a) classroom quality and teacher-child interactions, (b) teacher beliefs and practices, (c) teacher feelings about job stress, job control, job satisfaction, and depressive symptoms, and (d) teacher-rated child behavior problems for the two children in each classroom that the teacher identified at the pretest interval as most concerning.

EFFECTS ON CLASSROOM QUALITY AND TEACHER-CHILD INTERACTIONS

Classrooms referred for ECCP services appear similar in overall quality to the general population of state-funded prekindergarten classes in Connecticut. As presented in Tables 6 and 7, overall classroom quality as measured by the ECERS-R for both the treatment and control groups at posttest was similar to ECERS-R scores previously reported for Connecticut state-funded prekindergarten programs in south-central Connecticut ($M = 5.09$, $SD = 1.13$) (Gilliam, 2000).

However, no statistically significant treatment-control posttest differences were found for the ECERS-R total score or for any of the seven domains of quality measured by the ECERS-R in either Cohort 1 or Cohort 2. (See Tables 6, 7, and 8.) Similarly, no statistically significant treatment-control posttest differences were found for the three teacher-child interaction scales from the CIS measuring amount of positive interactions, degree of teacher detachedness, or degree of teacher punitiveness for either cohort. (See Tables 9, 10, and 11.)

EFFECTS ON TEACHER BELIEFS AND PRACTICES

Controlling for baseline differences at the pretest interval, no statistically significant differences were found for either cohort on the four scales measuring teachers' (a) beliefs about the importance of obedience, (b) beliefs about the importance of child independence, (c) beliefs about developmentally appropriate classroom practices, or (d) self-reported developmentally appropriate classroom practices they use in their classrooms. (See Tables 12, 13, and 14.)

EFFECTS ON TEACHER FEELINGS

Controlling for baseline differences at the pretest interval, no statistically significant differences for either cohort were found on any of the three subscales of teacher job stress (job control, job demands, or job satisfaction) or on the measure of self-reported depression. (See Tables, 15, 16, and 17.)

EFFECTS ON TEACHER-RATED CHILD BEHAVIOR PROBLEMS

Changes in the severity of behavior problems and amount of positive social skills, as reported by teachers, was measured for the two children in each classroom reported by their teacher to have the most concerning levels of behavior problems. Controlling for baseline differences at the pretest interval, statistically significant effects for ECCP were found on all scales of externalizing (acting-out) behavior problems on the two measures used in this evaluation. However, no effects were found for internalizing (shyness, withdrawn behaviors, anxiety, etc.) behaviors or positive social skills.

Across these measures, pretest scores for both the treatment and control groups showed clinically concerning levels of behavioral problems, indicating that the children targeted for ECCP services on average were indeed expressing severe levels of teacher-rated behavioral problems in their classrooms. At pretest, oppositional behaviors and hyperactivity across the two cohorts were on average in the area considered to be clinically significant (i.e., 1½ or more standard deviations above the norm or a T-Score of 65 or greater). (See Table 18, 19, and 20.)

On the CTRS, in the first cohort design where a randomized experiment was employed, a significant ($F_{(1,71)}=4.73$, $p<.05$, $d=0.53$) reduction for the treatment group relative to the control group in oppositional behaviors was noted. The severity of oppositional behaviors decreased for both the treatment and control groups from pretest to posttest, but the decrease for the treatment group was significantly greater than the decrease for the control group. The Cohen's d standardized effect size ($d=0.53$) was moderate. Effects were greater for hyperactivity ($F_{(1,71)}=6.83$, $p<.05$, $d=0.64$). Significant effects for ECCP was also noted on the composite scale for restlessness and impulsivity ($F_{(1,71)}=4.03$, $p<.05$, $d=0.48$). No significant effects were noted for any of the scales measuring internalizing behavior problems (e.g., anxious-shy, perfectionism) or general problems with social relatedness. Pretest measures for internalizing behaviors, however, were not on average in the clinically significant range, indicating that these behaviors typically were not very severe in the target children at the point of referral.

In the second cohort, a statistically significant ($F_{(1,67)}=5.77$, $p<.05$, $d=0.59$) positive effect was also found on the CTRS oppositional scale. However, the statistically significant Cohort 1 effects for hyperactivity and restlessness-impulsivity were not replicated in the second cohort. (See Table 19.) Combined data for both Cohorts are presented in Table 20.

On the SSRS, statistically significant effects in the first cohort were only found for total behavior problems ($F_{(1,71)}=3.95$, $p<.05$, $d=0.47$), where externalizing and internalizing behavior problems were combined. Although effect sizes for externalizing and internalizing behavior problems were nontrivial ($d=0.35$ and $d=0.39$, respectively), statistical power was not sufficient to detect these effects as being statistically significant. (See Table 21.) In the second cohort, effects for externalizing behavior problems approached statistical significance ($F_{(1,67)}=3.39$, $p=.07$, $d=0.45$). (See Table 22.) When data from both cohorts were combined, statistical power was sufficient to detect a small but positive effect in externalizing behavior problems ($F_{(1,141)}=5.28$, $p<.05$, $d=0.39$) and total behavior problems ($F_{(1,141)}=6.25$, $p<.05$, $d=0.42$). (See Table 23.) No statistically significant effects of ECCP were found, however, on any of the social skills scales of the SSRS for either cohort, with effect sizes in the trivial range (-0.15 to +0.10).

DISCUSSION

The effects of early childhood mental health consultation, as delivered by ECCP, were evaluated in a random-controlled crossover evaluation design. Overall, ECCP demonstrated statistically significant decreases in teacher-rated externalizing or acting-out behavior problems in the classroom, relative to the control group. The effects generally were consistent across the two measures employed, and were mostly of moderate magnitude. Effect sizes were greatest in the areas of oppositional behaviors and hyperactivity. There is little evidence of successfulness at reducing internalizing behavior problems (e.g., anxiousness, shyness, perfectionism, emotional lability). At pretest, however, specific children targeted for ECCP services did not on average evidence clinically significant levels of internalizing problems. Therefore, the ECCP intervention seemed to be more targeted toward children with externalizing behavior problems, and was more successful at reducing externalizing behaviors in the classroom. No evidence of effects on positive social skills (e.g., cooperation, self-control) was found. Whether the positive effects for reducing externalizing behavior problems would be observable to a trained, objective, rater who was blind to whether the classroom was receiving ECCP services, however, was not measured in this evaluation and remains unknown.

In contrast, no significant effects of ECCP were found for observable classroom quality, teacher-child interactions, teacher beliefs regarding discipline or classroom management, or teacher job stress and satisfaction. Using trained and reliable classroom raters who were blind to the treatment or control status of the classrooms they were observing, no significant effects of ECCP were found on the observational measures of global classroom quality or teacher-child interaction used in this evaluation, nor any of the subscales of these measures. Likewise, no significant differences were found for teachers' self-reported (a) beliefs about developmentally appropriate classroom practices, (b) developmentally appropriate practices they use in their classrooms, (c) beliefs about the importance of obedience, or (d) beliefs about the importance of child independence. Also, no significant differences were noted for teacher's job stress, sense of job control, job satisfaction, or depressive symptoms.

WHY WERE NO CLASSROOM OR TEACHER EFFECTS FOUND?

Given positive effects in child behavior, the lack of significant effects for classroom quality, teacher-child interactions, and teacher beliefs and feelings is puzzling. As previously discussed, the review by Brennan et al. (2005) reported some, albeit weak and inconsistent, evidence of ECMHC positive effects on classroom quality and teacher job stress and satisfaction. Moreover, Fink et al. (2003) noted that in ECCP's first year, 57% of teachers receiving ECCP services reported feeling that the quality of their classroom environments evidenced "great improvement." However, none of these findings were replicated in the present random-controlled evaluation of ECCP.

As a consultation service, ECCP is an indirect model of service. Consultative services were provided to the teachers, but no ECCP services were provided directly to children by the ECCP consultants. No evidence, however, was found to support any of the hypothesized pathways of effect (through improving classroom environments, changing teacher beliefs, or reducing teacher job stress and depression). Therefore, exactly *how* ECCP is effective at reducing teacher-rated externalizing behavior problems remains unknown.

Classroom effects may have been in areas too specific to be detected by these global classroom measures or may have been specific to individual children rather than the classroom environment on the whole. Perhaps the indirect pathways of effect from the consultant through the teacher and to the child are subtle, and the classroom and teacher measures were too global to detect effects in these areas. The classroom quality and teacher-child interaction measures used in the evaluation were not developed or validated as measures of consultation effect. Likewise, measures of teacher beliefs and feelings used in this study are global and may not be specific to the areas most likely addressed by a classroom mental health consultant. Therefore, many of the aspects of classroom quality, interactions, and teacher beliefs and feelings measured in this evaluation may not be the same as the foci of the consultative service.

Rather than effecting global classroom quality, perhaps ECCP was effective at changing teacher behaviors in more specific ways that were not adequately measured by these classroom instruments. Also, ECCP consultation focused on classroom-wide strategies as well as specific strategies for individual children with challenging classroom behaviors. However, the classroom quality and teacher-child interaction instruments measured change at the classroom level, whereas the child behavior instruments measured change at the individual child level. It is possible that classroom and teacher-child interaction effects may have been only at the individual child level for children targeted for specific ECCP service, and these effects may have been missed by the more global classroom measures used in the evaluation.

Perhaps differences in consultant focus and skills or the intensity of services may have diminished the overall effects in the area of classroom quality, teacher-child interaction, and teacher beliefs and feelings. Although the ECCP model is loosely manualized and there is a considerable amount of shared supervision that may increase consistency between consultants, ECCP consultants come from a variety of backgrounds with varying levels of training and experience. Also, average service intensity in this brief consultation model ranged considerably, from a low of 2.4 hours per week to a high of 6.5 hours per week. However, the level of variability in ECCP appears to be far less than that which is typical across early childhood mental health consultants nationally (Green & Everhart, 2006).

SIGNIFICANCE AND LIMITATIONS

Although much has been written about the provision of mental health consultation and related services in early education settings (Donahue et al., 2000; Perry et al., 2007; Hirokazu Yoshikawa & Knitzer, 1997), there is little published evidence of the effectiveness of mental health consultation that does not focus on specific manualized techniques. As such, this evaluation is the first random-controlled evaluation of a widely implemented system of early childhood mental health consultation (ECMHC). Although the pathways of effect remain unclear in this indirect model of consultation, ECCP was effective at improving teacher-rated externalizing behavior problems.

A limitation in this evaluation, however, is the sole reliance on teacher-ratings of children's behavior with no measures of child behavior collected by trained outside behavioral raters who are blind to treatment condition. It remains possible that the positive effects on children's behavior is based largely on teacher perceptions that may not be detectable to raters

who are unaware of whether the classroom had received ECCP services or not. Treatment condition blind raters were used for the classroom quality and teacher-child interaction measures, where no treatment effects were found. Further evaluations of ECMHC should use a variety of informants for child behavior impacts, including teacher report, parent/caregiver report, and classroom behavioral observations.

The ECCP model is well described, but it falls short of the manualized approaches of specific classroom behavioral techniques. Because the ECCP model is fluid and implementation may vary depending on both teacher and consultant input, this degree to which treatment fidelity can be assessed and managed is limited. Although flexibility of delivery may be advantageous in many respects, the lack of a clear manual and measures of treatment fidelity may create obstacles for exporting the ECCP model to other communities where ECCP administrative staff is not available. Interventions that have demonstrated effectiveness in both home and early childhood settings, such as Parent-Child Interaction Therapy (Filcheck, McNeil, Greco, & Bernard, 2004; McIntosh, Rizza, & Bliss, 2000; McNeil, Eyberg, Eisenstadt, Newcomb, & Funderburk, 1991), may provide a useful guide regarding the standardization and manualization of the intervention.

REFERENCES

- Alkon, A., Ramler, M., & MacLennan, K. (2003). Evaluation of mental health consultation in child care centers. *Early Childhood Education Journal*, 31, 91-99.
- Arnett, J. (1989). Caregivers in day care centers: Does training matter? *Journal of Applied Developmental Psychology*, 10, 541-552.
- Boyd, J., Barnett, W. S., Bodrova, E., Leong, D. J., Gomby, D., Robin, K. B., et al. (2005). *Promoting children's social and emotional development through preschool - Policy Report*. Retrieved May 27, 2006, from <http://nieer.org/resources/policyreports/report7.pdf>
- Brennan, E. M., Bradley, J. R., Allen, M. D., Perry, D. F., & Tsega, A. (2005, March). *The evidence base for mental health consultation in early childhood settings: Research synthesis and review*. Paper presented at the Establishing the Evidence Base for Early Childhood Mental Health, Tampa, FL.
- Cohen, J. (1988). *Statistical power analysis for the social sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Conners, C. K. (1997). *Conners' Teacher Rating Scale - Revised (L)*. North Tonawanda, NY: Multi-Health Systems.
- Cortina, J. M., & Nouri, H. (2000). *Effect size for ANOVA designs*. Thousand Oaks, CA: Sage.
- Curbow, B., Spratt, K., Ungaretti, A., McDonnell, K., & Breckler, S. (2000). Development of the Child Care Worker Job Stress Inventory. *Early Childhood Research Quarterly*, 15, 515-536.
- Donahue, P. J., Falk, B., & Provet, A. G. (2000). *Mental health consultation in early childhood*. Baltimore: Paul H. Brookes Publishing.
- Filcheck, H. A., McNeil, C. B., Greco, L. A., & Bernard, R. S. (2004). Using a whole-class token economy and coaching of teacher skills in a preschool classroom to manage disruptive behavior. *Psychology in the Schools*, 41, 351-361.
- Fink, D. B., Wakai, S., & Bruder, M. B. (2003). *Addressing behavioral outcomes in Connecticut's child care centers: An evaluation of the Early Childhood Consultation Partnership (ECCP)*. Farmington, CT: A. J. Pappanikou Center for Excellence in Developmental Disabilities, University of Connecticut Health Center.
- Gilliam, W. S. (2000). *The School Readiness Initiative in South-Central Connecticut: Classroom quality, teacher training and service provision*. Retrieved February 20, 2007, from nieer.org/resources/research/CSRI1999.pdf
- Gilliam, W. S. (2005). *Prekindergarteners left behind: Expulsion rates in state prekindergarten programs*. Retrieved May 27, 2006, from http://www.fcd-us.org/PDFs/NationalPreKExpulsionPaper03.02_new.pdf
- Gilliam, W. S., & Leiter, V. (2003). Evaluating early childhood programs: Improving quality and informing policy. *Zero to Three*, 23(6), 6-13.
- Gilliam, W. S., Ripple, C. H., Zigler, E. F., & Leiter, V. (2000). Evaluating child and family demonstration initiatives: Lessons from the Comprehensive Child Development Program. *Early Childhood Research Quarterly*, 15, 41-59.
- Gilliam, W. S., & Shahar, G. (2006). Preschool and child care expulsion and suspension: Rates and predictors in one state. *Infants and Young Children*, 19, 228-245.
- Green, B. L., & Everhart, M. (2006, February 24). *Multi-level determinants of effective mental health consultation: Results from a national survey*. Paper presented at the 19th Annual

- Research Conference, Research and Training Center for Children's Mental Health, Tampa, FL.
- Green, B. L., Simpson, J., Everhart, M. C., Vale, E., & Gettman, M. G. (2004). Understanding integrated mental health services in Head Start: Staff perspectives on mental health consultation. *NHSA Dialogue: A Research to Practice Journal for the Early Intervention Field*, 7(1), 35-60.
- Gresham, F. M., & Elliot, S. N. (1990). *Social Skills Rating System Manual*. Circle Pines, MN: American Guidance Service.
- Hamre, B. K., & Pianta, R. C. (2004). Self-reported depression in nonfamilial caregivers: Prevalence and associations with caregiver behavior in child-care settings. *Early Childhood Research Quarterly*, 19, 297-318.
- Harms, T., Clifford, R. M., & Cryer, D. (1998). *Early childhood environment rating scale (Rev. ed.)*. New York: Teachers College Press.
- Howes, R. B., Calkins, S. D., Anastopoulos, A. D., Keane, S. P., & Shelton, T. L. (2003). Regulatory contributors to children's kindergarten achievement. *Early Education and Development*, 14, 101-119.
- Keane, S. P., & Calkins, S. D. (2004). Predicting kindergarten peer social status from toddler and preschool behavior. *Journal of Abnormal Child Psychology*, 32, 409-423.
- Knitzer, J. (2000). *Using mental health strategies to move the early childhood agenda and promote school readiness*. New York: National Center for Children in Poverty, Columbia University.
- Knitzer, J. (2001). *Learning from Starting Points: Findings from the Starting Points Assessment Project*. New York: National Center for Children in Poverty, Columbia University.
- Marcon, R. A. (1999). Differential impact of preschool models on development and early learning of inner-city children: A three-cohort study. *Developmental Psychology*, 35, 358-375.
- McIntosh, D. E., Rizza, M. G., & Bliss, L. (2000). Implementing empirically supported interventions: Teacher-child interaction therapy. *Psychology in the Schools*, 37, 453-462.
- McNeil, C. B., Eyberg, S., Eisenstadt, T. H., Newcomb, K., & Funderburk, B. (1991). Parent-child interaction therapy with behavior problem children: Generalization of treatment effects to the school setting. *Journal of Clinical Child Psychology*, 20, 140-151.
- Perry, D. F., Kaufmann, R. K., & Knitzer, J. (2007). *Social and emotional health in early childhood: Building bridges between services and systems*. Baltimore: Paul H. Brookes Publishing.
- Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D., et al. (2005). Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions? *Applied Developmental Science*, 9, 144-159.
- Powell, D., Fixsen, D., & Dunlap, G. (2003). *Pathways to service utilization: A synthesis of evidence relevant to young children with challenging behaviors*. Tampa: Center for Evidence-Based Practice, University of South Florida.
- 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.
- Schaefer, E., & Edgerton, M. (1985). Parental and child correlates of parental modernity. In I. E. Sigel (Ed.), *Parental belief systems: The psychological consequences for children* (pp. 287-318). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Yoshikawa, H. (1995). Long-term effects of early childhood programs on social outcomes and delinquency. *The Future of Children*, 5(3), 51-75.
- Yoshikawa, H., & Knitzer, J. (1997). *Lessons from the field: Head Start mental health strategies to meet changing needs*. New York: National Center for Children in Poverty & American Orthopsychiatric Association.
- Zigler, E., Taussig, C., & Black, K. (1992). Early childhood intervention: A promising preventive for juvenile delinquency. *American Psychologist*, 47, 997-1006.

Table 1
Sample Sizes for Both Cohorts and Total

	Group Assignment	
	Treatment <i>n</i>	Control <i>n</i>
Cohort 1 (January 2005-June 2005)	23	23
Cohort 2 (September 2005-March 2006)	20	19
TOTAL	43	42

Table 2
Teacher Characteristics

	Cohort 1		Cohort 2	
	ECCP (<i>n</i> = 23)	Control (<i>n</i> = 23)	ECCP (<i>n</i> = 20)	Control (<i>n</i> = 19)
Gender				
Female	95.4%	100.0%	100.0%	88.9%
Race/Ethnicity ^a				
White (non Latino)	86.4%	81.8%	76.2%	68.4%
Black (non Latino)	0.0%	4.6%	9.5%	10.5%
Latino	9.1%	9.1%	9.5%	15.8%
Other	4.6%	4.6%	4.8%	5.3%
Holds a BA degree or higher	47.6%	77.2%	62.0%	64.7%
Age				
Mean	39.7	36.0	37.0	37.9
SD	12.2	11.2	11.7	9.4
Years Experience in Preschool				
Mean	9.0	8.3	7.5	11.0
SD	7.3	7.2	6.0	6.8
Class Enrollment				
Mean	17.0	15.7	15.2	19.6
SD	6.6	5.2	4.5	8.5
Number of Assistant Teachers				
Mean	2.0	1.7	1.9	2.1
SD	1.4	1.0	1.7	0.9

^a Note. Due to rounding, percentages may not necessarily sum to exactly 100%.

Table 3
Child Characteristics

	Cohort 1		Cohort 2	
	ECCP (<i>n</i> = 39)	Control (<i>n</i> = 35)	ECCP (<i>n</i> = 36)	Control (<i>n</i> = 34)
Gender				
Male	71.8%	77.1%	63.9%	76.5%
Female	28.2%	22.9%	36.1%	23.5%
Age in Years at Pretest				
Mean	4.09	4.31	4.04	4.03
SD	0.99	0.72	0.71	0.60

Table 4
Amount of ECCP Consultation Received Across Both Cohorts

	<i>n</i>	<u>Weeks of ECCP Consultation</u>					<u>Hours of ECCP Consultation</u>				
		Mean	SD	Median	Min	Max	M	SD	Median	Min	Max
Cohort 1	23	10.6	1.8	11	8	13	42.1	10.4	40	24	64
Cohort 2	20	11.3	1.2	12	9	13	44.1	9.7	44	30	61
Both	43	10.9	1.6	12	8	13	43.1	10.0	41	24	64

Table 5
Measures used across both cohorts for all four stages of data collection

	ECERS-R	CIS	PMS	JSI	CES-D	CTRS	SSRS
Cohort 1							
Pretest			X	X	X	X	X
Posttest	X	X	X	X	X	X	X
Cohort 2							
Pretest			X	X	X	X	X
Posttest	X	X	X	X	X	X	X

Table 6
Random-Controlled Experiment (Cohort 1) Posttest Differences on ECERS-R

	ECCP (<i>n</i> = 23)		Control (<i>n</i> = 23)		<i>F</i> _(1,44)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Space/Furnishings	5.13	1.07	4.84	0.97	0.96	0.29
Personal Care	4.75	1.25	4.25	1.23	1.92	0.41
Language/Reasoning	4.95	1.23	4.75	1.29	0.28	0.15
Activities	4.76	1.03	4.60	0.95	0.31	0.16
Interactions	4.89	1.33	4.79	1.54	0.06	0.07
Program Structure	5.31	1.24	5.54	1.16	0.43	-0.19
Parents/Staff	4.99	1.12	5.29	1.09	0.87	-0.28
ECERS-R Total	4.93	0.85	4.80	0.88	0.29	0.16

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 7
Quasi-Randomized Experiment (Cohort 2) Posttest Differences on ECERS-R

	ECCP (<i>n</i> = 20)		Control (<i>n</i> = 19)		<i>F</i> _(1,37)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Space/Furnishings	5.36	0.77	5.00	1.06	1.46	0.40
Personal Care	4.54	0.93	4.90	0.99	1.39	-0.38
Language/Reasoning	5.36	1.16	5.53	1.06	0.21	-0.15
Activities	5.10	1.02	4.82	1.06	0.70	0.27
Interactions	5.22	1.41	5.64	1.12	1.06	-0.33
Program Structure	5.71	1.29	5.51	1.43	0.21	0.15
Parents/Staff	4.74	0.70	5.24	0.97	3.30	-0.60
ECERS-R Total	5.11	0.71	5.15	0.79	0.02	-0.05

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 8
Combined (Cohort 1 + Cohort 2) Posttest Differences on ECERS-R

	ECCP (<i>n</i> = 43)		Control (<i>n</i> = 42)		<i>F</i> _(1,83)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Space/Furnishings	5.24	0.94	4.91	1.00	2.39	0.34
Personal Care	4.65	1.10	4.54	1.16	0.20	0.10
Language/Reasoning	5.14	1.20	5.10	1.24	0.02	0.03
Activities	4.92	1.02	4.70	1.00	1.00	0.22
Interactions	5.04	1.36	5.18	1.42	0.19	-0.10
Program Structure	5.50	1.26	5.53	1.27	0.01	-0.02
Parents/Staff	4.87	0.94	5.27	1.03	3.35	-0.41
ECERS-R Total	5.02	0.78	4.96	0.85	0.12	0.07

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 9***Random-Controlled Experiment (Cohort 1) Posttest Differences on CIS***

	ECCP (<i>n</i> = 23)		Control (<i>n</i> = 22)		<i>F</i> _(1,43)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Positive Interactions	2.94	0.45	3.03	0.57	0.34	-0.18
Not Overly Detached	2.11	0.30	2.03	0.29	0.71	0.28
Not Overly Punitive	1.53	0.45	1.39	0.28	1.57	0.37

* *p*<.05, ** *p*<.01, *** *p*<.001.**Table 10*****Quasi-Randomized Experiment (Cohort 2) Posttest Differences on CIS***

	ECCP (<i>n</i> = 20)		Control (<i>n</i> = 19)		<i>F</i> _(1,37)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Positive Interactions	2.93	0.41	2.89	0.31	0.12	0.11
Not Overly Detached	1.95	0.30	1.96	0.27	0.01	-0.04
Not Overly Punitive	1.28	0.22	1.41	0.35	2.07	-0.46

* *p*<.05, ** *p*<.01, *** *p*<.001.**Table 11*****Combined (Cohort 1 + Cohort 2) Posttest Differences on CIS***

	ECCP (<i>n</i> = 43)		Control (<i>n</i> = 41)		<i>F</i> _(1,82)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Positive Interactions	2.94	0.43	2.97	0.46	0.09	-0.07
Not Overly Detached	2.03	0.31	2.00	0.28	0.30	0.10
Not Overly Punitive	1.41	0.38	1.40	0.31	0.03	0.03

* *p*<.05, ** *p*<.01, *** *p*<.001.

Table 12
Random-Controlled Experiment (Cohort 1) Differences on PMS and BP

	ECCP (<i>n</i> = 20)		Control (<i>n</i> = 20)		<i>F</i> _(1,37)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
PMS-Obedience ^a						
Pretest	46.14	13.21	46.18	11.01		
Posttest	47.35	12.86	46.72	13.29	0.14	-0.13
PMS-Independence						
Pretest	35.68	3.87	34.25	3.09		
Posttest	34.18	6.34	33.50	3.50	0.00	0.00
Teacher Beliefs						
Pretest	54.31	11.36	49.25	7.60		
Posttest	53.56	10.55	48.44	7.99	0.23	0.16
Teacher Practices						
Pretest	49.67	13.51	47.44	6.17		
Posttest	50.59	10.29	46.94	7.36	0.61	0.26

^a Cohen's *d* is reversed for scales where a decrease in score is associated with an improvement.

* *p*<.05, ** *p*<.01, *** *p*<.001.

Table 13
Quasi-Randomized Experiment (Cohort 2) Differences on PMS and BP

	ECCP (<i>n</i> = 20)		Control (<i>n</i> = 19)		<i>F</i> _(1,36)	<i>D</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
PMS-Obedience ^a						
Pretest	47.25	13.45	50.47	11.73		
Posttest	49.76	12.79	49.79	11.63	1.16	-0.36
PMS-Independence						
Pretest	34.80	3.49	35.05	2.99		
Posttest	34.75	2.53	34.63	2.54	0.07	0.09
Teacher Beliefs						
Pretest	48.70	8.21	50.16	6.83		
Posttest	46.83	7.81	49.05	6.63	0.56	-0.25
Teacher Practices						
Pretest	46.05	8.25	46.66	6.86		
Posttest	45.22	6.26	46.61	5.85	0.53	-0.24

^a Cohen's *d* is reversed for scales where a decrease in score is associated with an improvement.

* *p*<.05, ** *p*<.01, *** *p*<.001.

Table 14
Combined (Cohort 1 + Cohort 2) Differences on PMS and BP

	ECCP (<i>n</i> = 40)		Control (<i>n</i> = 39)		<i>F</i> _(1,76)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
PMS-Obedience ^a						
Pretest	47.17	13.25	48.38	11.43		
Posttest	49.16	12.92	48.30	12.38	1.16	-0.25
PMS-Independence						
Pretest	34.58	4.53	34.94	3.19		
Posttest	34.27	4.65	34.32	3.15	0.02	-0.01
Teacher Beliefs						
Pretest	51.36	9.95	49.79	7.04		
Posttest	49.75	9.44	48.77	7.13	0.01	-0.03
Teacher Practices						
Pretest	47.16	6.41	47.15	10.99		
Posttest	46.82	6.46	47.10	9.04	0.05	-0.05

^a Cohen's *d* is reversed for scales where a decrease in score is associated with an improvement.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 15
Random-Controlled Experiment (Cohort 1) Differences on JSI and CES-D

	ECCP (<i>n</i> = 20)		Control (<i>n</i> = 20)		<i>F</i> _(1,37)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Job Control						
Pretest	52.67	9.27	53.92	11.65		
Posttest	49.24	10.05	52.89	8.94	1.54	-0.43
Job Demands ^a						
Pretest	54.27	14.93	46.58	11.32		
Posttest	55.29	16.83	51.17	9.98	1.75	0.47
Job Resources						
Pretest	73.21	6.39	69.95	10.43		
Posttest	71.40	8.29	67.70	11.38	0.22	0.15
CES-D ^a						
Pretest	8.68	7.24	6.10	4.62		
Posttest	8.95	10.59	10.07	7.16	1.56	0.43

^a Cohen's *d* is reversed for scales where a decrease in score is associated with an improvement.

* *p*<.05, ** *p*<.01, *** *p*<.001.

Table 16
Quasi-Randomized Experiment (Cohort 2) Differences on JSI and CES-D

	ECCP (<i>n</i> = 20)		Control (<i>n</i> = 19)		<i>F</i> _(1,36)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Job Control						
Pretest	51.15	7.29	54.53	10.29		
Posttest	50.80	5.80	55.11	11.09	0.90	-0.32
Job Demands ^a						
Pretest	51.50	11.57	55.84	16.14		
Posttest	51.10	12.23	52.11	16.11	1.05	-0.34
Job Resources						
Pretest	71.30	10.22	71.21	8.82		
Posttest	71.90	9.15	71.63	9.42	0.01	0.04
CES-D ^a						
Pretest	7.85	5.82	10.95	9.05		
Posttest	6.75	5.37	8.47	6.46	0.05	0.08

^a Cohen's *d* is reversed for scales where a decrease in score is associated with an improvement.

* *p*<.05, ** *p*<.01, *** *p*<.001.

Table 17
Combined (Cohort 1 + Cohort 2) Differences on JSI and CES-D

	ECCP (<i>n</i> = 40)		Control (<i>n</i> = 39)		<i>F</i> _(1,76)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Job Control						
Pretest	51.85	8.17	54.23	10.84		
Posttest	50.08	7.95	54.00	10.00	2.39	-0.36
Job Demands ^a						
Pretest	52.77	13.10	51.34	14.59		
Posttest	53.03	14.47	51.66	13.31	0.01	-0.02
Job Resources						
Pretest	72.25	8.47	70.56	9.57		
Posttest	71.65	8.62	69.61	10.52	0.23	0.11
CES-D ^a						
Pretest	8.26	6.48	8.88	7.46		
Posttest	7.82	8.29	9.25	6.76	0.54	0.17

^a Cohen's *d* is reversed for scales where a decrease in score is associated with an improvement.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 18
Random-Controlled Experiment (Cohort 1) Differences on CTRS

	Treatment (<i>n</i> = 39)		Control (<i>n</i> = 35)		<i>F</i> _(1,71)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Oppositional ^a						
Pretest	76.29	14.59	72.78	14.94		
Posttest	65.53	15.15	68.78	14.27	4.73 *	0.53
Hyperactivity ^a						
Pretest	72.13	12.17	67.84	12.63		
Posttest	64.62	14.04	67.42	10.88	6.83 *	0.64
Anxious-Shy ^a						
Pretest	61.32	13.33	58.42	11.32		
Posttest	58.82	13.01	56.03	9.35	0.26	-0.12
Perfectionism ^a						
Pretest	60.81	14.26	60.47	12.75		
Posttest	55.24	14.58	57.31	12.82	0.76	0.21
Social Problems ^a						
Pretest	67.86	15.70	65.22	16.07		
Posttest	66.84	18.67	64.31	15.63	0.02	-0.04
Restless-Impulsive ^a						
Pretest	70.31	12.95	67.33	12.52		
Posttest	63.41	14.22	65.48	11.51	4.03 *	0.48
Emotional Lability ^a						
Pretest	72.18	16.61	69.00	16.47		
Posttest	64.79	14.42	63.61	14.84	0.08	0.07
Total ^a						
Pretest	73.82	14.06	70.36	12.65		
Posttest	66.16	15.44	67.70	11.36	3.45	0.45

^a Cohen's *d* is reversed for scales where a decrease in score is associated with an improvement.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 19
Quasi-Randomized Experiment (Cohort 2) Differences on CTRS

	Treatment (<i>n</i> = 36)		Control (<i>n</i> = 34)		<i>F</i> _(1,67)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Oppositional ^a						
Pretest	68.11	17.49	68.88	15.71		
Posttest	60.94	16.08	67.59	15.60	5.77 *	0.59
Hyperactivity ^a						
Pretest	64.97	13.76	64.38	12.28		
Posttest	61.24	14.55	63.79	13.19	2.04	0.35
Anxious-Shy ^a						
Pretest	59.37	12.31	59.71	13.31		
Posttest	56.43	10.85	59.35	13.08	1.84	0.33
Perfectionism ^a						
Pretest	53.49	13.75	60.41	17.85		
Posttest	53.83	12.74	62.06	17.49	1.64	0.32
Social Problems ^a						
Pretest	66.46	15.99	65.85	14.82		
Posttest	65.63	17.74	62.62	14.08	0.80	-0.22
Restless-Impulsive ^a						
Pretest	62.51	14.01	64.29	13.28		
Posttest	60.29	13.74	62.88	11.74	0.46	0.17
Emotional Lability ^a						
Pretest	64.20	18.31	65.32	16.10		
Posttest	59.74	16.68	64.06	15.01	2.01	0.35
Total ^a						
Pretest	65.26	15.08	66.50	13.04		
Posttest	61.49	15.17	65.26	12.19	2.01	0.35

^a Cohen's *d* is reversed for scales where a decrease in score is associated with an improvement.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 20
Combined (Cohort 1 + Cohort 2) Differences on CTRS

	Treatment (<i>n</i> = 75)		Control (<i>n</i> = 69)		<i>F</i> _(1,141)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Oppositional ^a						
Pretest	72.37	16.46	70.77	15.35		
Posttest	63.33	15.67	68.17	14.86	10.96 **	0.57
Hyperactivity ^a						
Pretest	68.79	13.33	66.03	12.47		
Posttest	63.04	14.28	65.52	12.18	8.92 **	0.51
Anxious-Shy ^a						
Pretest	60.38	12.80	59.07	12.29		
Posttest	57.67	12.00	57.72	11.43	0.37	0.10
Perfectionism ^a						
Pretest	57.25	14.40	60.44	15.47		
Posttest	54.56	13.64	59.76	15.47	2.88	0.29
Social Problems ^a						
Pretest	67.18	15.74	65.55	15.32		
Posttest	66.25	18.11	63.44	14.76	0.61	-0.13
Restless-Impulsive ^a						
Pretest	66.62	13.93	65.79	12.90		
Posttest	61.93	13.99	64.16	11.61	4.10 *	0.34
Emotional Lability ^a						
Pretest	68.41	17.77	67.13	16.26		
Posttest	62.41	15.63	63.84	14.82	1.61	0.22
Total ^a						
Pretest	69.71	15.08	68.40	12.90		
Posttest	63.92	15.38	66.46	11.76	5.77 *	0.41

^a Cohen's *d* is reversed for scales where a decrease in score is associated with an improvement.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 21
Random-Controlled Experiment (Cohort 1) Differences on SSRS

	Treatment (<i>n</i> = 39)		Control (<i>n</i> = 35)		<i>F</i> _(1,71)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Cooperation						
Pretest	9.14	2.92	10.46	3.82		
Posttest	10.40	2.66	11.18	3.10	0.01	-0.03
Assertion						
Pretest	7.77	4.01	8.86	4.53		
Posttest	8.88	3.71	9.78	3.27	0.20	-0.11
Self-Control						
Pretest	6.61	3.35	8.15	3.80		
Posttest	8.34	3.35	9.33	2.87	0.29	-0.13
Total Social Skills						
Pretest	80.63	13.13	86.38	13.75		
Posttest	86.00	12.54	89.97	9.43	0.17	-0.10
Externalizing Prob ^a						
Pretest	8.23	3.06	7.73	2.62		
Posttest	6.26	3.63	6.69	2.81	2.18	0.35
Internalizing Prob ^a						
Pretest	2.26	2.22	1.91	1.80		
Posttest	1.41	1.68	1.80	1.95	2.74	0.39
Total Behavior Prob ^a						
Pretest	119.10	12.64	116.74	11.26		
Posttest	110.44	14.80	113.71	11.08	3.95 *	0.47

^a Cohen's *d* is reversed for scales where a decrease in score is associated with an improvement.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 22
Quasi-Randomized Experiment (Cohort 2) Differences on SSRS

	Treatment (<i>n</i> = 36)		Control (<i>n</i> = 34)		<i>F</i> _(1,67)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Cooperation						
Pretest	11.15	3.27	9.56	3.83		
Posttest	11.57	2.81	10.56	3.42	0.11	0.08
Assertion						
Pretest	8.61	3.68	7.15	4.03		
Posttest	9.39	3.56	9.15	3.80	0.38	-0.15
Self-Control						
Pretest	8.49	3.12	7.21	3.99		
Posttest	9.94	3.88	8.71	4.61	0.17	0.10
Total Social Skills						
Pretest	86.56	10.16	80.88	15.10		
Posttest	90.11	11.49	87.50	14.69	0.04	-0.05
Externalizing Prob ^a						
Pretest	6.33	3.36	7.15	3.17		
Posttest	5.03	3.29	6.62	3.13	3.39	0.45
Internalizing Prob ^a						
Pretest	2.08	1.61	1.71	1.61		
Posttest	1.86	1.66	1.71	1.68	0.04	0.05
Total Behavior Prob ^a						
Pretest	113.36	11.69	113.76	10.80		
Posttest	108.50	13.64	112.03	12.01	2.03	0.35

^a Cohen's *d* is reversed for scales where a decrease in score is associated with an improvement.

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 23
Combined (Cohort 1 + Cohort 2) Differences on SSRS

	Treatment (<i>n</i> = 75)		Control (<i>n</i> = 69)		<i>F</i> _(1,141)	<i>d</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>		
Cooperation						
Pretest	10.10	3.23	10.01	3.82		
Posttest	10.96	2.78	10.87	3.25	0.01	0.02
Assertion						
Pretest	8.18	3.85	7.99	4.34		
Posttest	9.13	3.62	9.46	3.53	0.81	-0.15
Self-Control						
Pretest	7.51	3.36	7.67	3.90		
Posttest	9.11	3.68	9.01	3.84	0.12	0.06
Total Social Skills						
Pretest	83.51	12.07	83.55	14.62		
Posttest	88.00	12.14	88.70	12.39	0.16	-0.07
Externalizing Prob ^a						
Pretest	7.32	3.33	7.44	2.90		
Posttest	5.67	3.50	6.65	2.95	5.28 *	0.39
Internalizing Prob ^a						
Pretest	2.17	1.94	1.81	1.70		
Posttest	1.63	1.67	1.75	1.81	1.82	0.23
Total Behavior Prob ^a						
Pretest	116.35	12.45	115.25	11.06		
Posttest	109.51	14.19	112.87	11.50	6.25 *	0.42

^a Cohen's *d* is reversed for scales where a decrease in score is associated with an improvement.

* $p < .05$, ** $p < .01$, *** $p < .001$.