Evaluation of the Illinois Model of Infant and Early Childhood Mental Health Consultation Pilot

Report to the Illinois Children's Mental Health Partnership

Julie Spielberger • Tiffany Burkhardt • Carolyn Winje • Aida Pacheco-Applegate • Elissa Gitlow • Erin Carreon • Lee Ann Huang • Anna Herriott • Reiko Kakuyama-Villaber

February 2021

CHAPIN HALL
AT THE UNIVERSITY OF CHICAGO

Recommended Citation

Spielberger, J., Burkhardt, T., Winje, C., Pacheco-Applegate, A., Gitlow, E., Carreon, E., Huang, L., Herriott, A., & Kakuyama-Villaber, R. (2021). *Evaluation of the Illinois Model of Infant and Early Childhood Mental Health Consultation Pilot*. Chicago, IL: Chapin Hall at the University of Chicago.

Acknowledgements

We sincerely thank the programs, supervisors, and teachers who participated in interviews and surveys and opened their classrooms to observation, and the families and home visitors who participated in surveys and interviews and allowed their home visits to be video-recorded. We also acknowledge other members of Chapin Hall who contributed to this evaluation—LaShaun Brooks, Leah Shapiro, Kristen Ethier, and Tracey Lockaby—and thank Matthew Brenner for his assistance in editing the report.

Past and current supporters of the evaluation: Anonymous funder, Caregiver Connections, Chicago Department of Family Support Services, Illinois Children's Healthcare Foundation, Irving Harris Foundation, McCormick Foundation, J.B. and M.K. Pritzker Family Foundation, W. Clement and Jessie V. Stone Foundation. This project was also funded in part by the Telligen Community Initiative to support, through research and programs, innovative and farsighted health-related projects aimed at improving the health, social well-being, and educational attainment of society, where such needs are expressed.

Public/private Leadership Team members: Illinois Action for Children, Caregiver Connections, Chapin Hall at the University of Chicago, Chicago Department of Family and Support Services, Chicago Department of Public Health, Cook County Health and Hospital Systems, Illinois Department of Children and Family Services, Illinois Department of Healthcare and Family Services, Illinois Department of Human Services, Illinois Department of Human Services-Division of Mental Health, Illinois Department of Human Services-Early Intervention, Illinois Department of Public Health, Erikson Institute, Governor's Office of Early Childhood Development, Illinois Association for Infant Mental Health, Illinois Chapter American Academy of Pediatrics, Illinois Children's Healthcare Foundation, Illinois Head Start Association, Illinois State Board of Education, Irving Harris Foundation, Lurie Children's Hospital, Start Early.

1313 East 60th Street
Chicago, IL 60637

chapinhall.org

Table of Contents

List of Tables

List of Figures

Executive Summary

Infant/Early Childhood Mental Health Consultation (IECMHC) is a relationship-based, collaborative support designed to improve the capacity of early childhood professionals to promote children's mental health (Cohen & Kaufmann, 2000). Social and emotional development is the foundation for success in learning and in life. It begins in children's early relationships with caregivers and is supported by strong partnerships among families, providers, programs, and systems. IECMHC consultants are trained mental health professionals who engage in a variety of promotion, prevention, and intervention activities to build the capacity of early childhood providers to foster social and emotional well-being and development of children and families.

A growing body of research has shown IECMHC to be an effective strategy for building the skills of early childhood staff and alleviating their job stress, as well as supporting children and families (Albritton et al., 2019; Brennan et al., 2008, Conners-Burrow et al., 2012; Duran et al., 2009; Hepburn et al., 2013; Perry et al., 2010). In addition, considering mounting evidence that preschool children of color experience harsh discipline at disproportionately higher rates than other children (Gilliam, 2005; Giordano, 2019; Giordano, et al., 2020; U.S. Department of Education Office for Civil Rights 2016), there has been strong interest in the potential of IECMHC to mitigate inequitable expulsions and suspensions from early childhood programs. The evidence base for IECMHC has evolved and expanded across multiple early childhood settings, including public pre-K, community-based childcare, and home visiting programs. However, efforts are often fragmented and lack a clearly defined model of consultation.

This report describes the results of a 3-year pilot study of the Illinois Model of IECMHC. The study was part of a 5-year comprehensive, coordinated, statewide initiative by The Illinois Children's Mental Health Partnership (ICMHP) to expand IECMHC across multiple systems and settings in Illinois. That initiative began in 2014, after almost two decades of coalition building and advocacy for IECMHC, when a private foundation convened public and private stakeholders to examine early childhood mental health in the state and develop a plan to integrate consultation into early childhood systems throughout Illinois (Harris Foundation, 2016).

The Illinois IECMHC Model

A broad-based Leadership Team of public and private stakeholders led the effort to develop the "Illinois Model" and provided oversight and guidance to the pilot implementation and evaluation. In addition to identifying the goals and critical elements of the model (see Box ES-1), the Leadership Team also established an infrastructure to embed IECMHC in multiple early childhood systems in the state for a sustained period. The infrastructure includes a common vision and funding commitment across diverse systems and communities and a workforce development strategy to ensure trained, highly qualified mental health consultants who can work across a range of settings.

All approaches to IECMHC aim to help to develop the skills of early childhood professionals to work more effectively with children and families. However, the Illinois Model is distinct in the priority it gives to relationship-building, reflective practice, and program-focused consultation as the means to build staff skills. (See Box ES-2 for types of consultation in the Illinois Model.) Relationships between consultants and staff are collaborative, ongoing, and proactive rather than episodic and reactive.

The theory of change for the Illinois Model assumes that if the approach is well-implemented and supported in multiple systems in diverse communities, then (1) administrators and staff will improve their reflective capacity, relationships with supervisors and coworkers, and knowledge of young children's and parents' social and emotional health; and (2) families and children will have more positive engagement with providers and easier access to high quality mental health services. In turn, (3) providers, families, and children will experience

Box ES-1. The Illinois Model

The Illinois Model of IECMHC is designed to be "universal," that is, applicable to a range of family- and child-serving systems and programs. In addition to identifying best practices, the model makes recommendations for coordinating consultation practices across the state and implementing the necessary structures and supports to ensure a high quality, diverse consultant workforce. It specifies a consultant's qualifications, competencies, and activities. The competencies are the following:

- Knowledge of infant/early childhood development, mental health, and early care and education
- Ability to build relationships and partner with families, providers, programs, and systems
- Ability to work effectively throughout diverse cultures and communities
- Ability to effectively and sensitively gather information
- Ability to collaboratively develop a plan and measures of success
- Knowledge of community systems and resources and ability to develop partnerships
- Commitment to ethical behavior and reflective practice

The Illinois Model is multi-level, flexible, and tailored to meet the needs and goals of the consultee(s). Thus, in practice, consultation can differ in its format, frequency and dosage, and focus or target. For the IECMHC pilot, consultants provided services 10-12 hours/month, on average, over 15 months, then 6 months of intermittent support. Activities were both program- and case-focused but prioritized relationships with staff and supervisors and building their knowledge and skills. Activities varied but included:

- Reflective consultation to individual staff or groups
- Support with observation, screening, and assessment of children
- Training on social and emotional development, the impact of trauma, or parental depression
- Co-facilitation of peer-support groups for program staff and/or caregivers
- Support for staff in meetings with parents

better outcomes. These outcomes include increased self-efficacy and reduced burnout and depression in staff and supervisors; positive social emotional development and better regulated behavior in children; and improved well-being and parenting practices in families.

The Illinois Model Pilot Study

The Illinois Model was piloted in three early childhood systems in four communities—Chicago urban, Chicago suburban, Peoria urban, and Peoria suburban/rural. The sample consisted of 23 early childhood programs, including center-based childcare and prekindergarten and home visiting programs. After matching, 15 programs were assigned to receive the Illinois Model of IECMHC.[1] The period of implementation varied somewhat between year-round programs and programs that closed during the summer. For year-round programs, the full implementation period was 21 months—15 months of intensive support and 6 months of intermittent support. The other eight programs, matched by type, served as a "business as usual" comparison group. Although some of the comparison programs received support from mental health consultants as part of their existing programs, none received services comparable to the Illinois Model.

Box ES-2. Types of Consultation in the Illinois Model

- ***Programmatic Consultation***: In collaboration with supervisors and directors, activities to assess and improve a program's structures, policies, procedures, professional development opportunities, philosophy, mission, and practices to better support the mental health of young children and families.
- ***Classroom and Home Consultation***: In collaboration with supervisors, staff, and parents, activities to assess and improve relationships, routines, and practices that affect the classroom or home climate.
- ***Child and Family Consultation***: In collaboration with families, staff, and other caregivers, activities to understand and respond effectively to an individual child's or family's mental health needs.

The goal of the evaluation was to assess both the implementation and effects of the Illinois Model pilot. The primary research questions were as follows:

(1) Was the Illinois Model of IECMHC implemented as intended? What factors affected its implementation?

(2) What were the effects of the intervention on staff and supervisors? Were there differences between staff in programs receiving the intervention and those in comparison programs in measured outcomes (reflective capacity, supervisor-staff relationships, burnout, depression, self-efficacy, and classroom and home visit environments)?

[1] Initially, there were 24 programs, 16 of which were in the intervention group, but one program withdrew from the pilot during the summer of 2019.

(3) What were the potential effects of the intervention on parent and child well-being and behavior? Were there differences between parents and children in programs receiving the intervention and those in comparison programs?

To address these questions, we used a mixed-methods, matched-comparison group design. During a three-month pre-implementation phase, mental health consultants were trained, and we collected baseline data. We collected additional data at three subsequent time points, approximately 6, 12, and 18 months after the start of implementation. The study drew from multiple sources of data: surveys and interviews with staff and supervisors, observations of classrooms and home visits, and assessments of children and families. Because the data were clustered and longitudinal, we analyzed the staff surveys, classroom observations, child assessments, and home visiting observations using linear mixed modeling, which accounts for missing data and the nonindependence of repeated measures in nested data (West et al., 2007).

Like most approaches to IECMHC, the Illinois Model both promoted use of specific strategies and had to be flexible and responsive to differences in programs in its implementation. Likewise, the evaluation design was rigorous but also had to be responsive to the community and program characteristics of the sample and variations in implementation. In addition to assessing the Illinois Model, this evaluation fills some important gaps in the literature. It provides more in-depth information about the process and challenges of implementing mental health consultation in early childhood systems, and a deeper understanding of the mechanism of change through which IECMHC impacts outcomes. In this report, we summarize our key findings by research question and discuss their implications for policy, practice, and further research.

Key Findings

Research Question 1: Was the Illinois Model of IECMHC implemented as intended? What factors affected its implementation?

Mental health consultants successfully implemented the Illinois Model based on structural and process indicators of fidelity. Despite a number of challenges with implementation, evaluation data on implementation dosage, adherence, and process indicate that the implementation of the Illinois Model was overall effective in both early childhood center-based programs and home visiting programs.

Structural indicators. We used two structural indicators to assess implementation. One was dosage, or the number of hours of consultation, and the other was adherence, or the extent to which consultants' activities were consistent with the model. Consultant logs indicated that all but two of the programs received at least 80% of their expected consultant goal hours. (One program did not because of a complete turnover in staff, and another had a structure that made it difficult for the consultant to meet as frequently with administrative staff as desired.) Although consultant activities varied, all intervention programs received the expected type of consultant

support. The most frequent activities were reflective supervision sessions with individual staff and their supervisors; reflective consultation with directors and supervisors; and reflective consultation with staff (without the supervisor present). There were differences between early childhood center-based programs and home visiting programs in types of activities. There also was considerable variability in activities among the programs in each group, reflecting the flexibility of the model to meet the characteristics and needs of individual programs.

Process indicators of fidelity. Qualitative interviews with program staff and consultants confirmed and added to the findings from the consultant logs. The interview data underscored the ways in which consultants adapted their work to fit the needs of the individual programs. While consultants spoke favorably of their training in the "Diversity-Informed Tenets for Work with Infants, Children, and Families" (Harris Foundation, 2016; Tenets Initiative, 2018), issues of diversity, equity, and inclusion were not a primary topic of consultation in most programs, reflecting an area for future growth in implementing IECMHC. The qualitative data also indicated similarities in the overarching needs of center-based early childhood and home visiting providers and how the Illinois Model can effectively support both types of programs.

Factors affecting implementation. As expected, it took time for consultants to build relationships with program supervisors and staff and develop processes for working together. Several factors impacted implementation. These included the ease or difficulty of scheduling meetings with staff and supervisors; stability or instability of staff at all levels (director, supervision, and staff); and extent to which leaders and staff understood IECMHC and their readiness to engage with the consultant. Indeed, one of the primary facilitators in successfully implementing the model was strong leadership support for consultation.

Research Question 2: What were the effects of the intervention on staff and supervisors? Were there differences between staff in programs receiving the intervention and those in comparison programs in measured outcomes?

Consistent with the theory of change for the Illinois Model, we found positive changes on two standardized measures of staff reflective capacity and a relationship between increased reflective capacity and decreased burnout in a subsample of staff. However, we did not see changes in standardized measures of staff-supervisor relationships (which were assessed quite favorably at baseline) or measures of burnout or depression (assessed low at baseline). Other factors, specifically, teacher position and race/ethnicity, appeared to have a stronger effect on these outcomes than the intervention did.

At the same time, there was evidence of an intervention effect on teachers' and home visitors' practices. Interview data confirmed this and revealed the following shifts in practice: 1) active listening and deeper exploration of issues; 2) the ability to think critically about one's reactions

and biases; 3) the ability to consider others' perspectives; and 4) the ability to establish or improve boundaries and be mindful of self-care.

Reflective capacity. Strengthening staff reflective capacity through reflective consultation is an important component of the Illinois Model of IECMHC. The intervention demonstrated positive effects on two measures of staff reflective capacity. The growth in staff reflective capacity was evident in both quantitative and qualitative data, whereas changes in supervisors were only apparent in the analysis of qualitative data, likely because of a small sample.

Being in the intervention group also significantly predicted lower emotional exhaustion, a component of burnout, at Time 3 for a subsample of staff, which was similar demographically to the larger sample. Growth in reflective process and collaboration predicted lower levels of emotional exhaustion, but the intervention was a stronger predictor. Thus, receiving the Illinois Model and building reflective capacity could mitigate staff burnout; however, we need additional research to better understand how consultation and improved reflective capacity can lead to lower burnout.

We also found group differences in burnout by race and ethnicity, indicating that these factors were bigger factors in burnout than the intervention. In particular, staff who identified themselves as White reported higher emotional exhaustion compared to all other racial and ethnic groups. Previous research has found that White providers tend to report higher burnout than Black and Hispanic providers (Salyers & Bond, 2001, in caseworkers; Garcia et al., 2020, in physicians). Although the reasons for these differences are unclear, it might reflect differences either in perceived burnout or in willingness to admit feelings of burnout.

We found that teacher role affected views of supervision and relationships with supervisors. Lead teachers in the intervention group had a more negative view of their supervisor's fidelity and delivery quality, efforts to build a bond or relationship with them, and efforts to support goals and tasks expected to benefit clients than lead teachers in the comparison group. One possible explanation for the difference is that after experiencing reflective conversations with the consultant, lead teachers in the intervention group realized that the supervision they received from their supervisor was not as reflective. Future research should further explore the effects of IECMHC on supervision and teachers' perceptions of supervision.

Teacher reflective capacity and child outcomes. Teachers with higher reflective capacity reported less teacher stress associated with children's behaviors; rated children's social and emotional strengths related to resilience greater; and rated children as having fewer problems with attention and emotion regulation than teachers with lower reflective capacity. Although directionality cannot be determined from these findings, strengthening reflective capacity might lead to lower teacher stress and shift teachers' perceptions of children to be more positive and strengths based. It is also possible that teachers' more positive views of children lead to less

stress and greater reflective capacity, as stress limits one's ability to be reflective (Ferguson, 2018). As Roffey (2012) noted, "How teachers feel makes a difference to their ability to respond effectively to the challenges they face" (p. 8).

Teacher depression and child outcomes. Although it is not clear whether IECMHC can affect measured depression in staff in the same way it can affect reflective capacity, depression is a variable that has been included in research on IECMHC (Silver & Zinsser, 2020). Greater reflective capacity was associated with teachers perceiving child behavior more positively, but teacher depression predicted more negative views of child behaviors and views of children's abilities to manage their behaviors. This association has a few possible explanations, as we cannot attribute causality: teacher depression could lead teachers to perceive child behavior more negatively; teacher depression could result in children exhibiting more behavioral concerns; or children's behavioral concerns and poor self-regulation skills could exacerbate teacher depression. Additional research could help to clarify this relationship.

Classroom climate. Observations in center-based classrooms showed that teachers in the intervention group were better able to manage children's behavior by enforcing clear, consistent, and developmentally appropriate rules of behavior and using proactive and positive behavior strategies over time than teachers in the comparison group. Teachers in the intervention group were also more likely to promote holistic development through a child-centered and individualized approach over time, although this finding was a trend that did not reach statistical significance.[2] These findings from the classroom observations suggest that center-based early childhood programs that received the intervention had a climate that better promoted mental health, particularly by responding to children in more positive, developmentally-appropriate ways, than programs who did not receive the intervention.

Equity in classrooms. Moreover, greater equity was observed in the classrooms of programs that were receiving the intervention than comparison programs. Diversity, equity, and inclusion is a core component of the Illinois Model. One core competency of the model is the consultant's ability to work effectively throughout diverse cultures and communities through cultural humility. These concepts were emphasized in consultant training before the initiative started and during the implementation of the model through ongoing training, supervision, and reflective learning opportunities, including workshops on the Diversity-Informed Tenets for Work with Infants, Children, & Families (Tenets Initiative, 2018). Thus, the finding that classrooms in the intervention group had higher ratings on equity is promising. However, it also underscores the

[2] Any finding reported as significant in this report is one with a *p* value of .05 or higher. Any finding reported as a trend or tendency is one with a *p*-value that approaches significance, i.e., is between .05 and .10.

need for further research on how the DEI core competency is reflected in home visiting and classroom practices and how to develop that competency.

Home visitor engagement. In the home visiting programs, we observed differences in the video-recorded observations of visits with staff who did and did not receive the intervention. Home visitors in the intervention group more frequently engaged in responsive behaviors during the home visit and elicited input on the content and activities of the home visit from parents than home visitors in the comparison group. In addition, there was a trend for home visitors in the intervention group to facilitate positive parent-child interactions and encourage the parent's leadership in the visit more often than home visitors in the comparison group. When we analyzed the home visit observation items that most aligned with the Illinois Model—essentially creating an IECMHC scale using the Home Visit Rating Scales-Adapted & Extended (HOVRS-A+; Roggman et al., 2010)—we found that home visitors who received the intervention tended to increase on this scale over time at a greater rate than those in the comparison group.

Research Question 3: What were the potential effects of the intervention on parent and child well-being and behavior? Were there differences between parents and children in programs receiving the intervention and those in comparison programs?

Child behavior. The evaluation did not assess children's behavior directly but relied on teachers' ratings. When teachers rated the severity of problems in children who they perceived to have behavioral problems on the Strengths & Difficulties Questionnaire, teachers in the intervention group tended to report less severe behavioral problems over time than teachers in the comparison group. Along with the classroom observation findings, this result supports the theory of change that mental health consultation for teachers can change both their practices to better support children's social and emotional development and their own perceptions of children's behavior. Because our measures were all teacher-reported, however, it is unclear whether these changes reflect actual change in children's behavior.

Contrary to some of the findings in the literature (e.g., Gilliam et al., 2016b), there were no racial or ethnic differences in teachers' assessments of children's behavior. However, consistent with the literature (e.g., LeBuffe & Naglieri, 2012), teachers rated girls significantly differently than boys on the child assessment measures, reporting more strengths and fewer challenges in girls compared to boys. The findings suggest that gender was the strongest influence on teachers' perceptions of children's behavior—stronger than race and stronger than mental health consultation. It may also suggest that another area of focus for mental health consultation is helping teachers better understand gender differences in children's development and behavior.

Family-level home visiting outcomes. Parents whose home visitors received the intervention tended to report higher satisfaction in their role as parents than parents whose home visitors

were in the comparison group. Features of the home visit were also associated with family-level outcomes. The home visitors' responsiveness to the family during home visits was associated with the parent's role satisfaction and parental report of a positive home environment. Home visitor practices to facilitate parent-child interactions were associated with the parent's report of the responsiveness and positivity in their interactions with their child. Consistent with the theory of change for the Illinois Model, home visitors' behaviors and aspects of the home visit predicted were also associated with parents' reports of positive interactions with their children.

Study Strengths and Limitations

This study makes important contributions to the growing body of IECMHC research literature. Several areas merit mention here.

- **Comprehensive, cross-system field study of both IECMHC implementation and outcomes**. This evaluation was the first to study the implementation of a new model of IECMHC in multiple early childhood systems, both school-based and community-based, using a matched-comparison group design. Although the variability in participating programs and consultants posed challenges for implementation, data collection, and analysis, it reflected the goal and commitment of the Leadership Team to examine implementation in the diverse communities and programs characteristic of Illinois.
- **IECMHC in home visiting**. One of the priorities of the Leadership Team was implementing the Illinois Model in home visiting programs. Few studies of IECMHC have included home visiting. We included six home visiting programs in the study and collected data from program supervisors, home visitors, and families, including recorded observations of home visits. We found positive effects of the intervention on home visitor practices. Specifically, home visitors who received IECMHC were more responsive to families and prioritized facilitating parent-child interactions during visits. In addition, implementation was somewhat easier in home visiting programs because their program structure included regular team meetings and supervision was more likely to incorporate reflection, in contrast to the program structure and supervision in early childhood center-based programs. On the other hand, home visiting programs in the study still experienced challenges in implementing the Illinois Model because of staff and director turnover and changes in funding and funders' requirements.
- **Innovative measures**. Many of the tools we used in this study were developed recently to measure constructs that are central to IECMHC but are also difficult to measure, such as reflective capacity and reflective supervision. First, to measure reflective capacity, we used the Provider Reflective Practice Assessment Scales (PRPAS; Heller, 2017). Although more research is needed to validate the tool, the PRPAS shows promise as a measure of change in reflective capacity. Second, we administered a standardized scale in the

surveys to measure reflective capacity, the Reflective Functioning Questionnaire (Fonagy et al., 2016). Third, we used the Reflective Supervision Rating Scale (Ash, 2010) to assess the content and structure of reflective supervision. For the classroom observations, we used the Climate of Healthy Interactions for Learning and Development (CHILD; Gilliam & Reyes, 2017), an observational assessment of the mental health climate of early care and education settings. The CHILD domains align very well with the aims and anticipated outcomes of IECMHC. Finally, based on the theory of change for the Illinois Model of IECMHC, the research team selected items from the HOVRS-A+ (Roggman et al., 2010) and created a new IECMHC scale for home visit observations.

- **Analytic approach**. We used linear mixed modeling (LMM) to account for the nested longitudinal data (e.g., children within classrooms within programs), missing values, and the many covariates. There were different numbers of staff and families per program, and the amount of time between data points was important to include, both of which LMM can address. Previous IECMHC evaluations that used a matched-comparison group design did not account for the clustered levels of the data (Conners-Burrow et al., 2012; Egeren et al., 2011; Gilliam, 2014). This is the first IECMHC evaluation to use both a matched-comparison group design and multilevel modeling.

As with every research study, our evaluation also had some limitations, which we discuss below. It is our hope that future research on IECMHC initiatives will consider these issues during the planning phase to ensure the strongest possible research designs.

- **Study timeline and scope**. Although the scope of the evaluation was a strength of the study, there were challenges associated with conducting both implementation and outcome studies at the same time. Ideally, an evaluation of a new model first should assess how the intervention is implemented and identify any barriers to implementation. An outcome study would occur only after there was evidence that the intervention or program was implemented as planned. This sequence would result in greater confidence that any observed outcomes could be attributed to a fully functional intervention, and any outcomes that were not observed were not due to implementation issues. However, we designed the evaluation to be responsive to the multiple information needs of the Leadership Team, prioritizing implementation and staff-level outcomes, but also examining the potential of the to affect children and families.

  There were not enough eligible programs in each setting, region, and community type to conduct a randomized control trial of the Illinois Model, which is typically considered the "gold standard" in evaluation design. We were able to use a matched-comparison group design to allow us to measure change that could be attributed to the intervention. Experts still consider the matched comparison group design to be a rigorous design

when it is not possible to randomly assign participants to study groups (e.g., Hanita et al., 2017). However, a limitation of this design is that we could not match programs on all potentially relevant program and staff characteristics before implementation started. The intervention and comparison groups were similar demographically at baseline, but they differed in staff education.

- **Comparison programs receiving consultation.** The programs in our comparison group were functioning as "business as usual," which means that they continued program operations as normal during the study. Several programs were receiving, or had access to, some form of mental health consultation during the study period. Although the consultation models were different from what the intervention programs were receiving, this may have masked measurable change of the Illinois Model on the intervention group in our analyses. We also lacked comprehensive information about the form and content of consultation in the comparison group, which limits our ability to explain differences or lack of differences in some of our outcome measures.
- **Variability in consultant relationships with programs.** There was considerable variability in the intervention programs' relationships with their mental health consultants. Some were familiar with the concept of mental health consultation or had an existing relationship with their assigned consultant prior to implementation, while others had never had a consultant before and had to develop relationships. As a result, the time it took to build relationships and trust between the consultant and the staff and the time to reach full implementation of the model varied across the programs. At the same time, this "limitation" also provided an opportunity to understand how the model will work once it is implemented more broadly.
- **Measure limitations.** Again, we selected a number of outcome measures developed over the past decade for use in evaluations of mental health consultation and related interventions. Although some measures have been used in diverse populations, one limitation is that others are still being tested and validated and may evolve further. Some measures do not have published psychometrics, and some might not have been sensitive enough to detect changes in staff and supervisor well-being and relationships that occurred because of the Illinois Model of IECMHC. For example, most staff reported low levels of burnout and positive relationships with supervisors, which meant that there was not a lot of room for improvement over time. Other researchers have suggested that baseline ratings may be artificially inflated, limiting ability to measure progress. For example, Heller and colleagues (2011) suggested that asking teachers to report on their own growth retrospectively after engaging in IECMHC might be more valid for some self-assessment measures than collecting self-report data at baseline.

- **Data collection challenges.** We collected data over three academic years, which caused some difficulty in terms of data quality and sample retention. Children moved classrooms and left programs. Some programs closed or had reduced programming during the summer; additionally, there was more turnover in program staff, including supervisors, than we were led to expect from the participating programs, which affected the ease of both program implementation and the evaluation. In particular, our sample of program supervisors was smaller than ideal, given how important the consultant-supervisor relationships are to the intervention.
- **Child assessments.** Unfortunately, we could not conduct the child assessments on a random sample of children. Instead, we asked teachers to select no more than eight children in their classroom whose parents had provided informed consent and who were likely to remain in the program the following year. Nonetheless, because the baseline data collection period was in the spring, transitions in staff and children during the summer resulted in a smaller sample of children who remained with the same teacher in the fall when the second data collection occurred.

Implications and Recommendations

This pilot study demonstrated several strengths of the Illinois Model. Establishing relationships and promoting infant and early childhood mental health through the parallel process (Johnston & Brinamen, 2006, 2012) are the foundation of the model. The model is preventive, aiming to support the well-being of children and families by building the capacity of the adults who care for and work with children, rather than only responding when challenges arise. The model uses reflective practice and a social justice framework to support and strengthen the early childhood care and education workforce. Its flexibility allows the approach to be implemented into different programs in different early childhood settings, each with its own set of challenges and needs. The study also resulted in several important findings relevant to practitioners, policymakers, and researchers interested in understanding what IECMHC can accomplish for program staff, families, and children. In the section below, we highlight some important considerations and implications of this research.

Practice Implications: The Illinois Model

Mental health professionals successfully implemented the Illinois Model in diverse settings, ranging from community-based childcare to school-based pre-K to home visiting programs. The consultants were well-trained and supported throughout the implementation, but they also varied in experience, understanding of the model, and prior relationships with the participating programs. Given all these variations, the model seems to have the right balance of structure and flexibility to be used in various settings by well-supported consultants from varied backgrounds. Implementation was facilitated by the infrastructure that was established by the Mental Health

Consultation Initiative, which encompassed more than this pilot study. Notably, the initiative has created a strong workforce development plan, started the development of a centralized data system, and obtained funding to continue to coordinate efforts to advance IECMHC across multiple early childhood systems.

Based on the results of the pilot study, our recommendations for the Illinois Model and its implementation fall into three main areas—program commitment to and readiness for implementation; flexibility of model; and workforce development, as follows.

Program Readiness and Commitment

- Ensure readiness of program staff to engage with consultant and establish structures for implementation. Complete a thorough readiness assessment prior to implementation to ensure all staff, not just directors and supervisors, understand the structure and process of the Illinois Model and are engaged from the beginning. Depending on their understanding, some programs might need more support to become ready to engage with the consultant. Indeed, the first several months of implementation might be labeled a readiness or preparatory phase of the Illinois Model.
- Establish minimum requirements and clear expectations for the consultation, including a regular schedule of meetings and space for the consultant.
- Continue to monitor implementation through data collection and periodic check-ins to make sure structures and schedules are working. Provide booster trainings every six months for staff and leadership in the model's approach or more often during times of staff transition.

Model Flexibility

- Maintain the flexibility of the Illinois Model's approach. Again, program administrators and staff will have varying levels of readiness, and some may need more support than others to fully engage with a consultant. Program structure, size, and staff needs will affect the monthly amount of consultation required. Our study findings suggest that 10-12 hours per month is appropriate for many larger programs, but smaller programs that do not have the schedules to allow for regular reflective supervision sessions may not have this much time. A consistent structure and schedule based on staff size might be more important than a specified number of hours. In addition, given the time it took some consultants to establish relationships with program staff at the beginning, more hours in the early months might help to solidify these relationships and ensure that staff and supervisors understand the Illinois Model's approach to consultation.
- Continue the consultation practices currently recommended by the Illinois Model while also monitoring their implementation to understand how they are working in different

programs. For example, the model advocates that consultants meet with staff and their supervisors together rather than individually. This helps to ensure good communication and relationships between supervisors and staff. Although some study participants, including a few consultants, resisted this idea at the beginning, over time they came to understand its value. Yet, some programs found it very difficult to coordinate schedules and put it into practice.

- Explore and be open to other means of communication with administrators and staff. The unfortunate arrival of the COVID-19 pandemic as the pilot was wrapping up forced some early childhood programs to experiment with the delivery of consultation services through virtual means.

Workforce Development

- Continue to monitor implementation with online data collection by consultants. Periodically share data with programs leaders and staff to help them understand the process and progress of regular consultation.
- Maintain ongoing supports and training for consultants. All consultants participating in this study appreciated the regular monthly supervision and ongoing opportunities to reflect and learn provided to them during the implementation. These supports were particularly important for less experienced consultants, with more seasoned consultants serving as mentors for less experienced consultants. Consultants highlighted the reflective learning groups, which provided regular opportunities to reflect with peers, as especially beneficial for a number of topics, for example, issues of diversity, equity, and inclusion (DEI).
- Relatedly, provide more in-depth training and support to help consultants implement the Diversity-Informed Tenets. This study found that consultants were familiar with and endorsed the Tenets because of training, but they varied in their skills and comfort in addressing them with program staff. Although our study did not focus in-depth on DEI, this area, which is so important to IECMHC, seemed difficult for many consultants to address. Consultants also reported that it was challenging to find the appropriate time and space for sensitive and uncomfortable conversations about DEI, particularly when program leaders did not recognize the relevance of these issues. These findings suggest a need for more intense training and, perhaps, more effective strategies and tools for consultants to use in implementing the Tenets, including how to initiate conversations related to DEI with program staff and administrators in order to support their growth in being culturally sensitive, aware, and humble.
- Try to match consultants and programs so that consultants have experience with the system in which they are working. We found that staff and supervisors appreciated

consultants who understood the content, funding requirements, and structures of the program they were serving.

Policy Implications

Illinois Inclusion Policy and IECMHC

The Illinois preschool expulsion ban legislation (Public Act 100-0105) was passed just prior to the start of the study. This law prohibits any program receiving funding from ISBE or licensed by DCFS from expelling children for behavioral reasons as of January 1, 2018. This legislation highlighted IECMHC as an important resource for staff in this legislation. If programs could no longer remove children, they need alternative solutions and resources to support them. This study adds to the growing body of evidence suggesting IECMHC is an effective support for early childhood program administrators and staff to develop new strategies for working with children who they perceive as having challenging behaviors.

Early Childhood Workforce

Research shows that young children and families benefit from high quality early childhood experiences, it is not easy for providers to achieve the level of quality necessary to support child development. The ability of early childhood center-based programs to meet the needs of children and their families depends, more than anything, on the professional development, knowledge, and skills of their staff. Over the past two decades, educational requirements for staff and program quality standards in publicly-funded programs, including Head Start, state pre-K, and home visiting, have become increasingly rigorous (Bernoteit et al. 2016), yet it has been difficult for the early childhood workforce to keep pace with new requirements. As a result, the workforce has widely varying qualifications, degrees, and credentials as well as compensation, which typically differ by funding stream.

Although IECMHC can support program staff facing these challenges, it is not enough to address all of the current issues and inequities in early childhood systems. IECMHC cannot be implemented successfully in a fragile system or fix systemic issues that contribute to staff stress, burnout, and turnover. For example, in community-based programs in this study, insufficient staff prevented consultants from facilitating reflective supervision because the supervisor had to serve as backup for a staff member. How can the Illinois Model make room to support programs with these kinds of barriers so that there is space for consultation rather than it feeling like an additional task on the list? Consultants showed themselves to be creative and adept at finding times to meet with supervisors and staff, but it was not easy. For IECMHC to be successful, staff must have time and space free of other responsibilities to meet with the consultant.

Research Implications

This study contributes to a growing body of research that has demonstrated positive effects of IECMHC for staff and families. However, we need additional research to determine whether the Illinois Model of IECMHC leads to reductions in disparities, as theorized, as well as longer-term outcomes such as staff professionalization, staff retention, improvements in behavioral regulation in children, and reductions in harsh disciplinary practices. We highlight some of our suggestions for additional research below.

- Conduct a follow-up study of program participants in this pilot to understand the sustained effects of consultation and structures put in place to keep consultation in place. The COVID-19 pandemic has changed service delivery, especially in school-based and home visiting programs. Evaluating the implementation and sustainability of IECMHC during challenging times like these is necessary. Early childhood programs likely need the support of mental health consultants now more than ever.
- Do more study of implementation, paying special attention to differences between programs in different early childhood systems to better understand adaptations that should be made for different program types and differences within the childcare or home visiting systems. There has been very little study of IECMHC in other early childhood systems such as family childcare, public health, and Early Intervention.[3]
- Explore the role of supervisors in IECMHC implementation and outcomes. Supervisors are less likely to be a focus of research on IECMHC but are integral to supporting the efforts of consultants to improve the knowledge and skills of frontline staff. Based on interviews with supervisors and consultants in this study, there was clear benefit for supervisors. However, our sample was very small, and standardized measures did not find differences between the two groups of supervisors. Thus, we recommend more study of the role of supervisors in implementing IECMHC, the challenges they experience in their work, and the supports they need to work more effectively with frontline staff.
- To better understand outcomes of the Illinois Model for children and families, conduct an experimental or quasi-experimental study of the model with a longer study timeline and larger sample of children; for example, a study that follows different cohorts over time as they transition to kindergarten. The child and family outcome data suggest that the Illinois Model has the potential to affect children and families in the long run but more rigorous, longitudinal studies are needed to understand its impacts. Furthermore, future research should measure the rates of child expulsion and suspension at the

[3] A small pilot study of the Illinois Model in four public health settings in Illinois is nearing completion but otherwise, we are not aware of other published research on IECMHC in public health settings.

program level, if possible, to determine any impacts IECMHC may have on preventing expulsions and suspensions of young children.

- Examine how mental health consultation can improve the equity of early childhood settings for diverse populations. Classroom observations and staff surveys in this study revealed some differences by staff race and ethnicity. For example, White teachers had lower scores on staff-child interactions and equity in their classrooms. It would be helpful to further analyze data from this study and other studies using the same classroom observation measure (CHILD) to examine the role of teacher-children racial concordance and discordance (i.e., same vs. different racial identity) on the classroom climate.
- Work with other researchers to develop more sensitive measures of the changes expected from IECMHC to more clearly assess the outcomes and mechanisms of change of consultation, including reflective practice, supervisor-staff relationships, staff well-being, and ability to promote children's and families' social and emotional growth. The measures of reflective capacity used in this study are very promising, although the PRPAS takes time to administer and analyze. Furthermore, we need psychometric evidence for some measures to ensure reliability and validity, especially for use in evaluations of IECMHC. Finally, our results suggested several relationships between variables, for example, reflective capacity and burnout, reflective capacity and perceptions of children's behavior, and effects of staff role and staff race on outcomes. These relationships are ripe for further investigation.

Conclusion

Given the variations in implementation and the size of the samples in this evaluation, we find the outcomes for staff, children, and families to be promising. At the same time, the extent of changes in some of the outcomes (notably, reflective capacity and classroom practices) indicates that there is room for further growth in staff, for example, in their reflective capacity and the social-emotional climate in classrooms. In addition, we need more study of outcomes, especially for supervisors, children, and families. We were impressed that any of the changes in child and family measures were significant or trending towards significance, given the fact that these are more distal outcomes than staff outcomes.

Introduction

> I think how it. . . ultimately helps families and children is that if I'm listening to the feedback from my consultant based on my frustrations then I adjust what I'm doing and I provide my teachers with more support and more understanding and more listening. . . . I think it helps them be better teachers in the classroom and then that helps the kids and their families. —*Supervisor*

Infant/Early Childhood Mental Health Consultation (IECMHC) is a relationship-based, collaborative support designed to improve the capacity of early childhood professionals to promote children's mental health (Cohen & Kaufmann, 2000). IECMHC recognizes that social and emotional development serves as the foundation for success in learning and in life. It acknowledges the importance of strong partnerships among families, providers, programs, systems, and IECMHC professionals. Consultants are trained mental health professionals who engage in a variety of promotion, prevention, and early intervention activities to build the capacity of early childhood providers to foster positive social and emotional well-being and development of children and families.[4] A consultant's activities are wide-ranging and may focus on programs, smaller group settings (classrooms, homes), and individual cases (a child or parent).

As discussed in this report, a growing body of research has shown IECMHC to be an effective strategy in supporting children, families, and staff; building the skills of early childhood providers; and alleviating provider stress. In recent years, in light of mounting evidence of Black preschoolers experiencing harsh discipline at disproportionately high rates compared to other children (Gilliam, 2005), there has been strong interest in the potential of IECMHC to mitigate expulsions and suspensions. The evidence base for IECMHC and consultation has evolved and expanded across multiple early childhood settings, including public pre-K, community-based childcare, and home visiting programs. However, efforts are often fragmented and lack a clearly defined model.

[4] Although there are similarities among IECMHC, coaching in the Pyramid Model, and behavioral interventionists, each is unique in its application. In IECMHC consultants use their knowledge and skills in mental health and early childhood to form relationships with and build the capacity of providers and other important adults in a child's life to help the child develop social/emotional and mental heath skills. A Pyramid Model coach focuses on ensuring program staff understand the Pyramid Model and are implementing it with fidelity, as well as responding when issues arise. Behavioral specialists focus on the needs of an individual child, develop a plan for a team to implement to address those needs, and monitor that plan and the child's progress. (Linda Delimata personal communication 12/14/2020).

This report describes the results of a 3-year pilot study of the Illinois Model of IECMHC. The study was part of a 5-year comprehensive, coordinated, statewide initiative by The Illinois Children's Mental Health Partnership (ICMHP) to expand IECMHC across multiple systems and settings in Illinois. That initiative began in 2014, after almost two decades of coalition building and advocacy for IECMHC, when a private foundation convened public and private stakeholders to examine early childhood mental health in the state. The coalition developed the *Plan to Integrate Early Childhood Mental Health into Child- and Family-Serving Systems, Prenatal through Age Five* (Harris Foundation, 2016).

A key goal of the plan was to implement a quality IECMHC approach to ensure that staff who work within any early childhood system can have regular access to reflective consultation and professional development about mental health issues, social and emotional development, and child and family well-being. A broad-based Leadership Team[5] of public and private stakeholders led the

Box 1. The Illinois Model

The Illinois Model of IECMHC is designed to be "universal," that is, applicable to a range of family- and child-serving systems and programs. In addition to identifying best practices, the model makes recommendations for coordinating consultation practices across the state and implementing the necessary structures and supports to ensure a high quality, diverse consultant workforce. It specifies a consultant's qualifications, competencies, and activities. The competencies are the following:

- Knowledge of infant/early childhood development, mental health, and early care and education
- Ability to build relationships and partner with families, providers, programs, and systems
- Ability to work effectively throughout diverse cultures and communities
- Ability to effectively and sensitively gather information
- Ability to collaboratively develop a plan and measures of success
- Knowledge of community systems and resources and ability to develop partnerships
- Commitment to ethical behavior and reflective practice

The Illinois Model is multilevel, flexible, and tailored to meet the needs and goals of the consultee(s). Thus, in practice, consultation can differ in its format, frequency and dosage, and focus or target. For the IECMHC pilot, consultants provided services 10-12 hours/month, on average, over 15 months, then 6 months of intermittent support. Activities were both program- and case-focused but prioritized relationships with staff and supervisors and building their knowledge and skills. Activities varied but included:

- Reflective consultation to individual staff or groups
- Support with observation, screening, and assessment of children
- Training on social and emotional development, the impact of trauma, or parental depression •
- Co-facilitation of peer-support groups for program staff and/or caregivers
- Support for staff in meetings with parents

[5] An organization chart for the Infant/Early Childhood Mental Health Consultation Initiative is included in Figure A-1 in Appendix A.

effort to develop the "Illinois Model" and provided oversight and guidance to the pilot implementation and evaluation. Box 1 summarizes the goals and critical elements of the model.

The Case for IECMHC

Children's relationships and social-emotional development are the foundation of their success in learning and in life. This is the fundamental assumption of IECMHC, a support to build the capacity and skills of early childhood professionals who care for and work with young children and their families. The RAINE Group (2014) summarized how IECMHC can address several current problems in the early childhood workforce. For example, about one-third of early childhood teachers report high levels of job stress and burnout, which can lead to poorer classroom climates (Jennings, 2015), as well as higher rates of child expulsion and staff turnover (Whitebook & Sakai, 2003). Up to one-fifth of young children in child care have serious social-emotional problems, and many more children have less severe, but still challenging, behaviors that not only interfere with their learning and development, but also interfere with the teacher's ability to provide instruction to other children (Davis & Perry, 2014; Egger & Angold, 2006; Gilliam, 2005). In addition, home visiting program staff have identified several training needs in topics related to mental health that IECMHC addresses. For example, Korfmacher et al. (2012) found that almost half (46%) of home visitors wanted training in working with caregivers with significant challenges, such as depression, substance abuse, and domestic violence. In addition, 44% wanted training in working with families whose children have serious behavioral or mental health concerns.

IECMHC supports an existing program or service rather than being a distinct or direct service to children or parents. As such, it is difficult to distinguish its discrete effects from the effects of the program itself. However, previous experimental and quasi-experimental studies have compared providers and children in group settings with and without the support of mental health consultation or have looked for changes before and after a consultation (e.g., Gilliam et al., 2016a; Hepburn et al., 2013). Research indicates that mental health consultation in early care and education programs increases teachers' ability to understand and respond appropriately to children's behaviors and needs, reduces teacher stress, and increases job satisfaction (Albritton et al., 2019; Brennan et al., 2008). However, as discussed below, we know much less about the effects of consultation on providers in other child- and family-serving organizations (Albritton et al., 2019), such as home visiting programs, and even less published evidence exists about the impacts on parents and children in these other settings.

Process Evaluations of IECMHC Implementation

There has been considerable research on implementing consultation, such as whether the intervention was implemented as planned (i.e., fidelity; Carroll et al., 2007) and served the intended population, and the duration required for the intervention to be fully implemented and

change to occur (Hepburn et al., 2013). Studying how an initiative or program was implemented is critical for interpreting the results of an outcome evaluation and understanding the context in which outcomes were achieved. If a program or intervention was not delivered as intended, changes in the intended outcomes cannot be expected. We have included recent research on these aspects of IECMHC implementation in this section.

Fidelity. Implementation evaluations of IECMHC have tried to examine fidelity by assessing whether the program delivered aligns with the model or approach planned to be implemented. Because IECMHC is considered a flexible model that has used different approaches, assessing fidelity is not straightforward. Some programs have used a customized, flexible approach to consultation, combining an evidence-based curriculum (e.g., the Incredible Years; see Webster-Stratton, 2005) with individualized IECMHC, such as the model tested by Raver et al. (2009). Other IECMHC methods of implementation have used a manualized approach to consultation. Manualized models, however, still require evaluators to document adaptations made to ensure that the interventions fit within local contexts and for different populations (Perry et al., 2010). The Healthy Futures evaluation in Washington, DC highlighted the need to document the consultation protocols and develop an implementation manual to assist in training new staff and scaling up the intervention (Perry, 2013).

The process evaluation of Early Childhood Consultation Partnership (ECCP) in Connecticut (Fink et al., 2003; Gilliam, 2014) found both child- and classroom-specific services were delivered with moderately strong fidelity to the group of children targeted by the intervention. However, the number of children who received the child-specific services was far lower than anticipated—less than half of the expected number of children received intensive, child-specific treatments. Thus, the services were not delivered to the target population as planned.

Duration and dosage. The duration of the various IECMHC interventions to date has varied, ranging from 8 weeks (Gilliam, 2007; 2014) to 4 years (Shivers, 2015). For example, although Gilliam (2014) concluded that a longer duration of the intervention may be necessary to see significant change in classroom climate, they did find changes in child externalizing behavior after only eight weeks. The process evaluation of Arizona's Smart Support (Shivers, 2015) found that during the 12-month intervention period during which data were collected from each program, significant change occurred in the first 6 months of the intervention and was then sustained over the second 6-month period.

Factors that affect implementation. According to Duran et al. (2009) and others, the success of IECMHC depends in part on the readiness of early childhood programs, staff, and parents and families. Indicators of early childhood program readiness include the presence of a supportive early childhood program administrator/director, flexibility to incorporate consultation into the program, and embrace of a "mental health perspective" (i.e., a recognition that infant and young

children's behavior must be understood within the context of development, relationships, and how the environment impacts relationships; Cohen & Kaufmann, 2000). Ash (2009) developed an instrument to measure a program's readiness to engage in IECMHC, The Kid Connects Early Care Site Readiness Assessment tool.[6] Indicators of early childhood provider readiness are an openness to gaining more knowledge, a desire to try something new, a willingness to collaborate, and not feeling threatened by the consultant's involvement. Indicators of parent/caregiver readiness include an acceptance that there are issues negatively impacting the child that need to be addressed, a willingness to try something new, and a willingness to collaborate (Duran et al., 2009). Another component of readiness that could be useful to measure is implementation climate, a global construct consisting of items related to expectations, support, and rewards (Jacobs et al., 2014).

These implementation evaluations all occurred in center-based early childhood settings, yet IECMHC has been implemented in home visiting programs as well. A recent implementation evaluation of IECMHC in Colorado Maternal, Infant, and Early Childhood Home Visiting (MIECHV) programs examined the role of the consultant, how consultation was being implemented in this setting, and the barriers and facilitators to implementation (Colorado Department of Public Health & Environment, Early Childhood Evaluation Unit, 2020). The evaluation findings include, for example, that home visitor emotional and personal discomfort was a barrier to implementation, but an open environment and trust between home visitors and consultants facilitated implementation. Trusting relationships appear to be key to successful IECMHC implementation.

Although several IECMHC implementation evaluations have been conducted, more rigorous process evaluations are necessary to understand the effectiveness of implementation and fidelity of future IECMHC models. Measuring fidelity in programs that are intended to be flexible, like the Illinois Model, is both difficult and important. Without implementation evaluations, outcome evaluations cannot be properly conducted and interpreted.

Outcome Evaluations of IECMHC

Over the past two decades, the growing body of research on IECMHC has focused largely on center-based early childhood programs. Reviews of multiple studies of the effectiveness of IECMHC include those conducted by Brennan and colleagues (2008), Duran et al. (2009), Perry et al. (2010), and most recently, by Albritton et al. (2019). In an issue of Zero to Three dedicated to IECMHC, Hepburn, et al. (2013) discuss the status of evidence and briefly summarize seven

[6] The Kid Connects Early Care Site Readiness Assessment tool, designed and used in Colorado, includes the following rating categories: 1) Administrative Support for Mental Health Consultation, 2) Current Practices, 3) Flexibility of Programming to Incorporate Mental Health Consultation, and 4) Readiness for Partnership with Mental Health Consultation (Ash, 2009).

statewide IECMHC evaluations (Arizona, Arkansas, Connecticut, Washington, DC, Louisiana, Maryland, and Michigan). An external evaluation was also conducted on the statewide IECMHC in Pennsylvania (Davis & Perry, 2016). Findings from these and other evaluations show that consultation has the potential to affect multiple outcome levels, including classroom and program climate, provider and staff well-being and practices, and child and parent well-being. The literature also highlights the potential impact IECMHC can have on inequitable preschool suspension and expulsion practices (Davis, et al. 2020; Gilliam et al., 2016b; Hepburn et al., 2013; Silver & Zinsser, 2020). However, only three studies used experimental or quasi-experimental designs to evaluate IECMHC (Conners-Burrow et al., 2012; Egeren et al., 2011; Gilliam, 2014). This highlights the need for additional rigorous evidence demonstrating the effects of IECMHC.

Provider Outcomes

Because IECMHC affects providers most directly, many IECMHC evaluations have focused on provider-level outcomes. To date, most of the literature has focused on provider well-being (i.e., stress, burnout, depression, and self-efficacy), provider knowledge and skill (including provider–child interactions), and the implicit biases providers bring to their work. Indeed, a fundamental assumption of the Illinois Model is that in order to work effectively with children and families and support their well-being and growth, providers must not only be knowledgeable and skilled, but also mentally healthy. Providers' well-being can influence their perceptions of and reactions to children (Buettner et al., 2016; Jeon et al., 2019; Roffey, 2012).

Provider well-being. Given that teacher job stress has been found to be a strong predictor of children's expulsion rates (Gilliam & Shahar, 2006; Zinsser et al., 2019), job stress levels have been examined as an outcome of IECMHC. Alarmingly, early childhood teachers are experiencing increasing levels of stress and burnout (Jennings et al., 2020). Teacher burnout has been found to be associated with feelings of hopelessness, irritability, and impatience, substance abuse, absenteeism, increased turnover, and decreased job performance (Lowenstein, 1991). In some studies, teachers who have received IECMHC reported lower levels of stress and burnout (Brennan et al., 2008; Hepburn et al., 2013). Other aspects of teachers' mental health may be supported by IECMHC and may also influence expulsion rates. Teachers' depression has been associated with expulsion requests (Silver & Zinsser, 2020). IECMHC has partially moderated this association, suggesting that consultation may be a factor in decreasing expulsion rates (Silver & Zinsser, 2020).

Provider knowledge and skill. In studies by Brennan et al. (2008) and Hepburn et al. (2013), providers who have received IECMHC have reported increased self-efficacy and competence in dealing with difficult behaviors of young children in their care, as well as working with families. They found that not only did teachers feel more confident in their ability to support the social-emotional needs of children, they demonstrated an improved capacity for working with children.

In a Michigan study, for example, childcare providers reported an improved ability to recognize early signs of developmental, social-emotional, and behavioral concerns and improved interactions with children because of a higher dosage of IECMHC services (Hepburn et al., 2013). In Arizona, teachers receiving IECMHC demonstrated increased closeness and decreased conflict in their interactions with children (Shivers, 2015). In Arkansas, teacher–child interactions showed reductions in levels of punitiveness, permissiveness, and detachment, with a trend towards higher sensitivity following IECMHC (Conners-Burrow et al., 2012). In multiple studies, teachers also improved in their interactions that supported social and emotional development and showed increased teaching about feelings and emotional problem-solving skills (Hepburn et al., 2013). Classroom observations revealed increased emotional support by teachers receiving IECMHC (Hepburn et al., 2013). While increased positivity in teacher–child interactions was found in many IECMHC evaluations, Gilliam (2007) used observer ratings of teacher–child interactions but did not find differences between intervention and control groups.

In home visiting programs, some provider-level outcomes have been found as well. In two evaluations of IECMHC in home visiting programs, the majority of home visitors reported increased knowledge of children's social and emotional development and child behavior in context (Center for Prevention Research and Development, 2011; Goodson et al., 2013). However, both studies collected only self-report retrospective data on home visitors, asking them whether they gained knowledge in a number of areas. More rigorous research on the effects of IECMHC in home visiting programs is needed.

Implicit bias. Similarly, and following Gilliam (2005), Gilliam et al. (2016a), and others (Albritton et al., 2019; Zinsser, 2019), in this study we consider expulsion or expulsion risk as a classroom or provider outcome rather than a child outcome, while acknowledging the harm it can have on a child. As Meek and Gilliam noted, "Expulsions and suspensions are not child behaviors; they are adult decisions" (2016, p. 6). Examining the role that IECMHC may play in mitigating expulsions and suspensions is increasingly urgent as Black preschool children are suspended at disproportionately high rates (Meek & Gilliam 2016; U.S. Department of Education Office for Civil Rights, 2016). Research on preschool expulsions has examined the impact of IECMHC on expulsion rates, and Gilliam and Shahar (2006) found an association between IECMHC and reduced rates of expulsions, which was replicated in Washington, D.C. and Maryland (Hepburn et al., 2013).

Implicit bias likely plays a central role in teachers' expulsion requests and is believed to be a mechanism through which IECMHC could affect change (Davis et al., 2020; Meek & Gilliam, 2016; Shivers et al., 2018). Researchers have hypothesized that the increase in reflective capacity provided by IECMHC may be one pathway by which implicit bias is decreased (Davis et al., 2020), although there are few studies to date that have looked at the role of reflective capacity in mediating change.

Child Outcomes

Previous research evaluating the effects of IECMHC in early childhood programs has found multiple child-level outcomes. As IECMHC is intended to foster children's social-emotional development and skills (Duran et al., 2009), outcome evaluations have often measured social-emotional development and internalizing and externalizing behaviors. A main finding reported across studies is the reduction in children's externalizing behavior (Hepburn et al., 2013; Perry et al., 2010). For example, Gilliam (2007) found that effect sizes were greatest in decreased oppositional behavior and hyperactivity. In a family-centered model of consultation, the dosage of consultation predicted improved child behavior (Upshur et al., 2009).

In addition to measuring externalizing behavior, many studies have also measured internalizing behavior, such as withdrawal and depression, and prosocial behavior. The studies that have reported findings on internalizing behavior found improvements for children who received IECMHC (Bleecker & Sherwood, 2004; Green et al., 2006, Raver et al., 2009). Improvement in children's prosocial behavior is consistently reported in IECMHC research, including social skills, communication, social interactions, cooperation, self-control, play and leisure time, coping skills, interpersonal relationships, initiative, and attachment (Hepburn et al., 2013; Perry et al., 2010).

Parent/Family Outcomes

Although parents of children in early childhood education settings are not typically the direct recipients of IECMHC, some studies of IECMHC evaluations have measured its impact on parents. In studies that measured this, some improvements were found. The evaluation of Michigan's Childcare Expulsion Prevention Initiative (CCEP) found that parents whose children were enrolled in CCEP programs receiving IECMHC services showed increased empowerment in advocating for their children relative to the comparison group (Egeren et al., 2011). Parents reported reduced work/school problems after receiving IECMHC; only 18% of parents in the CCEP group had work/school problems at follow-up, while 100% of comparison parents did. Parenting stress, however, was not impacted by IECMHC, as both the intervention and comparison groups showed similar improvements in parenting stress. Furthermore, dosage of consultation was not associated with improvement in parenting stress and empowerment (Egeren et al., 2011).

In home visiting programs, parents are typically more directly targeted by IECMHC programs. Over half of mothers in home visiting programs struggle with issues related to mental health, domestic violence, and substance misuse (Tandon et al., 2005)., Families that face these problems were more likely to disengage or drop out of home visiting programs (Daro et al., 2003). Home visitors may find it difficult to identify mental health challenges and provide support to mothers with these struggles, especially if they have not received training in this area

(Tandon et al., 2005). Furthermore, home visitors have reported needing support in reflecting on and coping with this challenging work (Parlakian, 2001).

Classroom/Program Outcomes

Evaluations of IECMHC have found positive outcomes at the classroom level. Multiple studies have found the level of overall problems in the classroom was reduced (for review, see Hepburn et al., 2013). For example, in Maryland, teachers' perceptions of the rates of problem behaviors in their classrooms decreased significantly after receiving 4 months of onsite consultation. In addition, some studies found improvements in classroom organization (Louisiana & Washington, DC). In Gilliam's (2014) evaluation, most teachers reported "great improvement" in the quality of their classroom environments, activities and interactions. However, a later evaluation (Gilliam, Maupin, & Reyes, 2016) did not find significant differences in classroom climates between two groups of programs: one that received IECMHC and one that did not. Raver and colleagues (2008) found a post-intervention increase in positive classroom climate, decreased negative classroom climate, increased teacher sensitivity, and improved teacher classroom management.

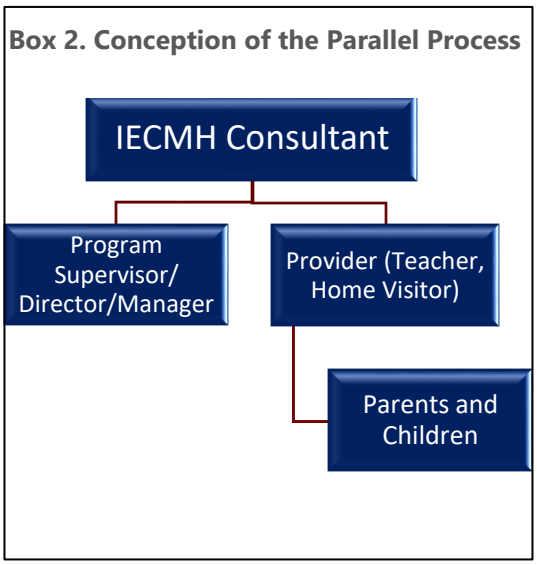
Overall, studies that assessed classroom environment using the Early Childhood Environment Rating Scale (ECERS) found mixed results, possibly because this tool is not sufficiently sensitive to measure the types of change targeted by IECMHC. Gilliam (2008) developed the Preschool Mental Health Climate Scale (PMHCS) in response to a lack of sensitivity in tools used in previous studies. Evaluations that used the PMHCS found strong positive results. For example, interactions were related to classroom quality in evaluations conducted in Washington, DC, Maryland, and Arizona. Shivers (2011) found that improvements in the classroom climate as measured by the PMHCS were associated with multiple IECMHC activities, including more time spent with a teacher, more time spent modeling interactions, more time spent observing children, and more written action plans developed for children.

A number of IECMHC evaluations report program-level outcomes. For instance, program quality increased after receiving IECMHC in five of the seven evaluations described by Brennan and colleagues (2008). Alkon, Ramler, and MacLennan (2003) found that the duration of the consultation predicted program quality: programs that implemented IECMHC for longer than one year showed significant improvement in childcare quality compared to programs that had consultation services for less than 1 year. The extent to which a consultant is integrated into program functioning is another factor that appears to influence the outcome of consultation. Program staff view consultation services as more effective when consultants are well-integrated and highly involved in the program and considered "part of the team" (Gilliam, 2005; Green, et al., 2004). Finally, programs receiving IECMHC services tended to have reductions in staff

turnover (Brennan et al., 2008). Research indicates that lower levels of staff burnout and stress correspond with lower levels of staff turnover (Brennan et al., 2008; Hepburn et al., 2013).

Mechanism of Change

The positive child and provider outcomes shown through research on IECMHC evaluations raise the question: What is the mechanism by which IECMHC achieves these outcomes? Although there is no definitive answer to this question, the literature suggests one potential mechanism of change is the relationship between the consultant and the early childhood provider. Green et al. (2006) found that "the single most important characteristic of mental health consultants is their ability to build positive collaborative relationships with program staff members." Positive relationships between the consultant and early childhood care and education staff have been labeled a "catalyst for success" for positive child, family, staff, and program outcomes (Duran et al., 2009) in that these relationships are expected to change relationships between program staff and the children or parents with whom they work. Johnston and Brinamen (2006, 2012) describe the parallel process at work in consultation, explaining that the transformative power of the consultation–provider relationship allows the provider to develop new ways of interacting with children and other adults (see Box 2).



Shivers (2015) found that consultant–teacher relationships moderated the effect of the intervention on child-level and teacher–child relationship outcomes. When Green and colleagues (2006) conducted a national survey about Head Start consultants, they found that the strongest predictor of perceived effectiveness of consultation was the quality of the relationship of the staff member with the consultant. Moreover, the quality of the consultant–provider relationship was associated with staff wellness (Green et al., 2006). Consultant characteristics and the amount of consultation time also have been shown to influence consultant–staff relationships (Duran et al., 2009). Furthermore, providers who reported strong relationships with consultants were more likely to report goals related to improved child and family well-being than staff who reported weaker relationships with consultants (Green et al., 2004). According to Allen and Green (2012), consultant reports of positive relationships with families, high levels of supervision and support, and consultant reports of positive relationships with staff are all positively associated with staff reports of positive relationships with the consultant. Finally, integration of the consultant into the program

and clearly delineated roles for consultants predicted a strong consultant-provider relationship (Green et al., 2006).

However, research that has attempted to identify the mechanism for change in child outcomes has yielded mixed results. Gilliam (2007) found changes in child outcomes but no teacher or classroom effects. Since consultation was not provided directly to children and no data were collected from consultants, there was no evidence supporting any of their hypothesized mechanisms of change. Allen and Green (2012) predicted that consultants' attributes would impact child outcomes, but none of the five consultant-reported attributes—knowledge of early childhood settings, relationships with families, cultural sensitivity, relationships with staff, and knowledge of IECMHC best practices—were significant predictors of improvements in children's internalizing or externalizing behaviors. Another example is the Arkansas IECMHC evaluation in which found changes in child outcomes only after year 3. The authors posit that it is possible that children did not change their behavior as much as teachers changed in their perceptions of the children's behavior (Conners-Burrow et al., 2012). Without the data to support these theories, however, the mechanisms through which the outcomes occur remain unclear.

These findings provide some insight into the processes by which IECMHC affects change. Yet, as highlighted by Perry et al. (2010), there is a need for increased rigor in evaluations, including independent assessments of child behavior. In some studies, the consultants who delivered the intervention also collected the data (e.g., Shivers, 2015), which limits the data's validity. Again, many studies measured child behavior by teacher report, raising the question of whether child *behavior* changed or teacher *perceptions of child behavior* changed, suggesting that their perceptions could be a mechanism by which children's behavior changes (Conners-Burrow et al., 2012; Gilliam et al., 2016a; Perry et al., 2010). Brennan and her colleagues (2005) noted that there are few reliable and valid tools available to measure the pathways through which mental health consultation may affect children's behaviors, such as the quality of the relationship between the teacher and the child or the teacher's internal representation of the child. Standardized measures of teacher/provider–child interactions do exist, e.g., the Arnett Caregiver Interaction Scale (CIS; Arnett, 1989) and the Student-Teacher Relationship Scale (STRS; Pianta, 2001); however, there are concerns about the psychometric properties of these measures and the lack of sensitivity to variations in caregiver skills or interactions with children (Colwell et al., 2013).

Many of the previous studies on IECMHC did not measure variables that may be important in mediating or moderating the outcomes of interest. Shivers (2011) found that in Arizona's Smart Start program, improvements in the classroom climate were associated with multiple consultant activities, including more time spent with the teacher, more time spent modeling interactions, more time spent observing children, and more written action plans developed for children. However, consultants' activities have not been documented consistently in IECMHC evaluations. The quality of supervision received by the consultants may be a moderating variable, although

supervision received by consultants was only reported in five of the 14 studies in Perry et al. (2010). Program location is another potential moderating variable that is rarely measured. One study found that program location led to differences between consultants' relationships with parents and the providers' relationships with the consultants: the effect of the consultant–parent relationship on the consultant–staff relationship was greater in urban programs than rural (Allen & Green, 2012). Perry and colleagues (2010) call for designs that test whether changes in children's outcomes are mediated by gains in teachers' skills or practices.

In summary, the benefits of IECMHC are supported by a large body of research evidence, particularly in early childhood center-based programs. The research to date indicates that mental health consultation can increase early childhood teachers' ability to understand and respond appropriately to children's problem behaviors, reduce their stress, and increase their job satisfaction, among other outcomes. However, we know less about the effects of mental health consultation on staff in other child- and family-serving settings, such as home visiting and Early Intervention Part C programs—or the impacts on parents or children in these settings. Only a few published descriptions of guidelines for effective implementation in home-based programs exist (Goodson et al., 2013; Pittenger et al., 2015).[7] Most notably, we are not aware of any literature on comprehensive consultation models that span multiple early childhood systems as the Illinois Model does. Finally, the research suggests limitations in the designs and the measures used in previous research on IECMHC that make it difficult to understand the process by which IECMHC affects children and families.

The Illinois IECMHC Model

The IECMHC model for Illinois is based on both the research reviewed above and on lessons learned from a decade of experiences with consultation in a range of early childhood settings in Illinois, including Head Start, preschool, childcare, and home visiting programs. The ICMHP Leadership Team determined that effective IECMHC implementation requires three core elements: strong infrastructure, highly qualified mental health consultants, and high-quality services. Members of the Leadership Team also recognized that there are many different approaches to delivering mental health consultation. However, based on experience and research, they believed that to increase the capacity of early childhood systems and maintain that capacity, the Illinois Model has to be embedded within programs with services provided on an ongoing, regular basis, as opposed to on an intermittent basis when a problem arises.

Toward that end, the Leadership Team developed a plan that established an infrastructure to embed IECMHC in multiple early childhood service systems in the state over a sustained time

[7] Another resource is an unpublished report prepared by the Center for Prevention Research & Development (2011) on ECMHC in Healthy Families Illinois and Parents Too Soon (PTS) home visiting programs.

period. The infrastructure includes a common vision and funding commitment across diverse systems and communities. It also includes a workforce development strategy to ensure there are trained, highly qualified mental health consultants who can work across a range of settings to develop positive relationships with staff and provide knowledge and services to enhance the skills of providers. The underlying assumptions are that a highly qualified, trained consultant who can develop rapport and facilitate reflection and collaborative problem solving with providers can improve their capacity to understand and respond appropriately to children and parents. Improvements in relationships with children and parents, in turn, are expected to increase their engagement in and ability to obtain the benefits of early childhood programs and services.

The Illinois Model Pilot Structure

The Illinois Model for the pilot study included a training and orientation phase for participating systems, followed by 21 months of ongoing services from trained Master's level professionals with mental health, clinical, and educational backgrounds. Consultants were expected to have the necessary knowledge, skills, and experiences to serve staff and programs across multiple systems. Consultants provided approximately 10 to 12 hours of training, consultation, and referrals to program staff each month during the first 15 months, and fewer hours (an average of 2 hours/month) during a 6-month intermittent support phase. They received training in the Illinois Model prior to implementation as well as ongoing training and supervision from the pilot implementation team during the implementation period.[8] The implementation team also developed a set of printed materials, including a menu of potential activities, for the pilot study.

However, the actual activities of the consultants were partly a function of the individual needs of home visitors, teachers, classrooms, or programs.

[8] The training infrastructure plan is outlined in a recent (2017) Workforce Development Plan developed by a committee of the ICMHP Leadership Team. The goal of the plan is to build and sustain the I/ECHMHC workforce ability to serve diverse systems through the standardization of orientation and training opportunities, the expansion of reflective learning groups, coaching/mentoring support, and other professional development opportunities, including support for consultant credentialing and certification.

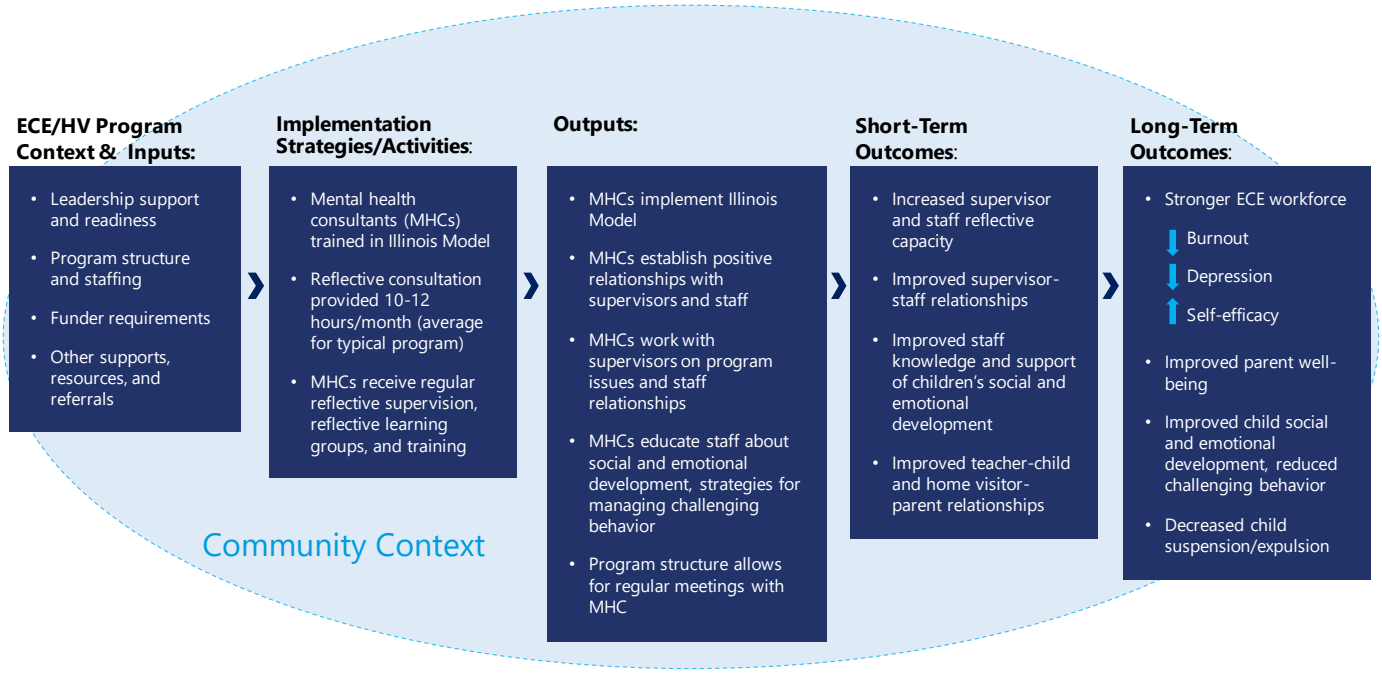
During the first year, consultants were expected to meet with the on-site program leader or supervisor(s) every month to provide reflective consultation. They were also expected to meet the agency or program administrator at least twice a year. In addition, consultants were to meet with all staff in groups at least once a month to provide training, case consultation, and reflective supervision, and to consult with individual staff and supervisor together twice a month. Again, other site-specific planned activities (e.g., observations of home visits or classroom activities) depended on the needs of particular programs and staff. The actual focus of the consultant's activities also varied based on staff needs and ranged from program-focused to child-focused consultation. (See Box 3.)

Box 3. Types of Consultation in the Illinois Model

- ***Programmatic Consultation***: In collaboration with supervisors and directors, activities to assess and improve a program's structures, policies, procedures, professional development opportunities, philosophy, mission, and practices to better support the mental health of young children and families.
- ***Classroom and Home Consultation***: In collaboration with supervisors, staff, and parents, activities to assess and improve relationships, routines, and practices that affect the classroom or home climate.
- ***Child and Family Consultation***: In collaboration with families, staff, and other caregivers, activities to understand and respond effectively to an individual child's or family's mental health needs.

All approaches to IECMHC help develop early childhood professionals' skills while also strategizing around challenging behaviors in children and families. However, the Illinois Model is distinct in the priority it gives to relationship building, reflective practice, and program-focused consultation as the means to build staff skills: "The consultant works to create a safe opportunity for individuals to communicate and reflect on aspects of the system, program, practices and themselves to develop an understanding and awareness that strengthens their capacity to support children" (Illinois Model of IECMHC Model Overview). Relationships between consultants and staff are collaborative, ongoing, long-term, and proactive rather than episodic and reactive. As displayed in Figure 1, the Illinois Model theory of change assumes that if the approach is well-implemented and supported in multiple systems in diverse communities, then (1) administrators and staff will improve their reflective capacity, relationships with supervisors and coworkers, and knowledge of young children's and parents' social and emotional health; and (2) families and children will have more positive engagement with providers and easier access to high-quality mental health services. In turn, (3) providers, families, and children will experience better outcomes. These outcomes include reduced burnout and depression and increased self-efficacy in staff and supervisors; positive social emotional development and better regulated behavior in children; and improved well-being and parenting practices in families.

Figure 1. Theory of Change for the Illinois IECMHC Model



Another central feature of the Illinois Model to note is its focus on diversity, equity, and inclusion. A core competency of the Illinois Model is the consultant's ability to work effectively throughout diverse cultures and communities through cultural humility. Diversity, equity, and inclusion were emphasized in consultant training before the initiative started and during the implementation of the model through ongoing training, supervision, and reflective learning opportunities, including training on the Diversity-Informed Tenets for Work with Infants, Children, and Families (Tenets Initiative, 2018) The Tenets include Diversity-Informed Practice, Equity, Infant and Early Childhood Mental Health, Intersectionality, Privilege, Reflective Capacity, and Social Justice. According to Shivers and her colleagues (2018), early childhood mental health consultants are viewed as champions of equity and disrupters of implicit bias.

The Pilot Study of the Illinois Model

The Illinois Model was piloted in early childhood systems in four different communities: Chicago urban, Chicago suburban, Peoria urban, and Peoria suburban/rural. The sample included 23 early childhood center-based and home visiting programs, 15 of which received the Illinois model of IECMHC as the Intervention group.[9] A matched group of eight programs were in a comparison

[9] Initially, there were 24 programs, 16 of which were in the intervention group, but one program withdrew from the pilot during the summer of 2019.

group that did not receive the model.[10] Among the programs in the intervention group, there was some variability in length of implementation. For year-round programs, the full implementation period was 21 months—15 months of intensive support and 6 months of intermittent support. For programs that closed for 2–3 months during the summer, the implementation period was somewhat shorter.

The goal of the pilot study was to evaluate the implementation and effects of the Illinois Model across child- and family-serving systems in the state. In addition to being an examination of the Illinois Model, this evaluation provides more in-depth information about the processes and challenges of implementing mental health consultation in early childhood systems and a deeper understanding of the mechanism of change through which IECMHC impacts outcomes. Key research questions were as follows:

(1) Was the Illinois Model of IECMHC implemented as intended? What factors affected its implementation?

(2) What were the effects of the intervention on staff and supervisors? Were there differences between staff in programs receiving the intervention and those in comparison programs in measured outcomes (reflective capacity, supervisor–staff relationships, burnout, depression, self-efficacy, and classroom and home visit environments)?

(3) What were the potential effects of the intervention on parent and child well-being and behavior? Were there differences between parents and children in programs receiving the intervention and those in comparison programs?

To address these questions, we used a mixed-methods, matched-comparison group design, which is described in more detail in the next chapter. We matched eligible early childhood center-based and home visiting programs, and then randomly assigned them to either an intervention group of 16 programs or a comparison group of eight programs. During a 3-month pre-implementation phase, mental health consultants were trained and we collected baseline data. We collected additional data at three subsequent time points, approximately 6, 12, and 18 months after the start of implementation. The study drew from multiple sources of data: surveys and interviews with program staff and supervisors, observations of classroom and home environments, and assessments of children and families.

[10] These eight programs conducted "business as usual." They did not have access to the Illinois Model, but some did receive periodic support from a mental health consultant, estimated to be 1 or 2 hours per month, on average.

Overview of this Report

In the next chapter, we describe the characteristics of the communities and programs that participated in the study and our research design and methodology. We present our findings in the subsequent two chapters. First, we present findings about the implementation of the Illinois Model, including structural indicators (such as the number of hours of consultation program staff received) and process indicators (such as the types of activities and issues addressed in consultation) to see if they were in line with the expectations of the Model. Next, we describe findings about the effects of the intervention on staff functioning, staff interactions with children and families, and child and family functioning. The final chapter summarizes all of the findings and discusses their implications for future practice, policy, and research.

It is important to note that we collected data for this study between April 2018 and March 2020, just before the COVID-19 pandemic forced the immediate closure of many early childhood programs in Illinois. As we are preparing this report, our study participants—families, program staff, and communities—are experiencing unprecedented levels of unemployment and health crises that disproportionately affect low-income communities. The pandemic has illuminated racial inequities in the structures of health care and other policies and institutions in our society. Given the timing of our study, the findings do not reflect the current context or what early childhood systems will look like in the future. Even so, we believe that the key results and recommendations remain current, and are ever more critical during these times.

Methods and Sample

> I feel like this has given us the opportunity to really listen to one another. It's helped us as a team, an entire team, with my assistant and the lead teacher next door and her assistant. I feel it's helped us work together a little more. I would keep it, and I would keep it for both new and experienced teachers. —*Teacher*

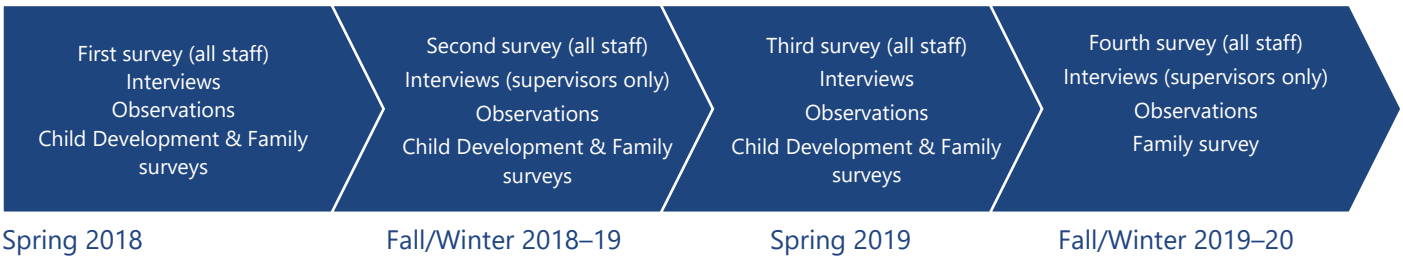
The Illinois Model of IECMHC is based on a theory of change and specifies the types of consultant activities that adhere best to the model's principles and assumptions. At the same time, like most approaches to IECMHC, it has to be flexible. The way consultation is delivered depends on the needs and goals of the program staff, the relationship between the consultant and staff, and a number of other factors, including organizational context and how early childhood staff understand and use consultants' support. We recognized, given the unique characteristics of each program, that the implementation—the consultants' roles and activities and the length of time it would take to build relationships with staff—would vary from program to program. Thus, we understood that the evaluation required a design that would address both implementation and outcome questions while being adaptive to changing circumstances.

This chapter describes the design, sample selection, and methods for the evaluation. In order to examine implementation processes and participant outcomes together, it was necessary to collect and triangulate both quantitative and qualitative data from a variety of participants and to look at change over the course of the 21-month implementation. We were interested in staff perspectives on the intervention, so we administered surveys and conducted interviews. We also wanted to collect data from other perspectives and sources. We collected observational data of teachers' interactions with children in their classrooms and home visitors' interactions with families. We also used data from consultants' logs of activities. Together these sources provide a rich picture of the implementation of the Illinois Model, its impacts on program staff and supervisors, and its potential to affect children and families.

Design and Timeline

The design of the evaluation for the pilot was developed in collaboration with the Leadership Team. The design took into account budget, funders' grant cycles, and the availability of qualified consultants. For example, training the consultants and collecting baseline data had to start in the spring to allow time to collect data and schedule initial visits with the programs before the start of summer, when a number of programs either closed or operated on a modified schedule. Figure 2 outlines the evaluation's timeline.

Figure 2. Study Timeline and Methods for Mental Health Consultation Initiative Pilot



Community Selection

The Leadership Team sought to conduct the pilot in four communities that represented the demographic and geographic diversity of communities in the state. The process of selecting the communities and, subsequently, the programs for the pilot ended up being complicated, as described below. Resource constraints and the availability of communities with a sufficient number of early childhood settings contributed to complicating the process.

The Leadership Team wanted to select four different communities to represent the diversity of communities in the state of Illinois. Toward that end, the initiative's pilot Implementation Team collected descriptive data about communities across Illinois. This information included total population; child and family demographic data; and the number of communities with an early childhood coalition or network. To be considered for the pilot, communities had to have multiple programs with a sufficient number of staff and program participants to ensure the sample size would meet the needs of the evaluation. Minimally, each prospective community was required to have one home visiting program, one childcare program, and one preschool program.

After analyzing the data, the Implementation Team identified twelve communities as eligible for participation. In consultation with the Leadership Team, six communities were prioritized based on the number of early childhood programs of different types, geographic area, and the extent to which its population represented the diversity of the state. Final decisions were based on proximity to allow for easier implementation, decreased travel time, and the availability and cost of consultants. Given these considerations, Chicago and its surrounding south suburbs in Cook County were two of the programs chosen. The other two were in the town of Peoria and surrounding rural towns in Peoria County, located in north central Illinois.

The selection of the communities occurred in the fall of 2017. Table 1 displays the characteristics of the selected counties according to 2017 U.S. Census data (Early Childhood Asset Map, 2020; U.S. Census Bureau, 2020).

Table 1. Characteristics of Selected Counties

Characteristics		Peoria County	Cook County
Population		183,011	5,150,233
Race/Ethnicity	Non-Hispanic White	70%	65%
	Black	19%	28%
	Hispanic/Latino	5%	24%
Income	Below Federal Poverty Level (State: 13%)	15%	14%
	Median household income (State: $59,196)	$51,632	$62,088

Program Selection and Recruitment

The Leadership Team sought to enroll 24 programs: 16 to receive the intervention and eight to serve as a comparison group. Members of two subgroups of the Leadership Team, the Implementation and Evaluation teams, met weekly over 2 months to make the final selection. Eligibility criteria for programs included being licensed and in compliance with their funders, monitors, or other applicable regulatory bodies. The goal was to select programs that met a threshold for a minimum number of staff (i.e., at least three staff members for home visiting sites and six staff members for early childhood center-based programs).

All potential programs in the selected communities were contacted by the state or city agency or organization aligned with their funding, including Illinois State Board of Education (ISBE), Illinois Action for Children, and the Chicago Department of Family Support Services (DFSS). The Implementation Pilot Director and, frequently, a representative of the broader initiative or the evaluation team visited each interested program to discuss its potential for participation in the pilot. During these visits, the Implementation Team Director explained in detail the implementation activities and responsibilities for participating in the evaluation as either an intervention or a comparison program. These meetings occurred prior to group assignment. Each program was also given written descriptions of the pilot and research activities. In addition, all participating programs were expected to engage in the evaluation process, send at least one supervisor or director to a pilot orientation meeting, participate in a regional orientation meeting with staff, and, if identified as an intervention program, set aside time to meet regularly with the consultant.

The goal of the pilot was to compare the Illinois Model of IECMHC with standard "business as usual" services, which meant either not receiving any IECMHC services or receiving some IECMHC services but at a much lower rate of frequency and intensity. Overall, program leaders were eager to participate in the pilot and enthusiastic about both potentially receiving IECMHC and adding to the knowledge base regarding IECMHC. To assist in our decision making, we collected extensive information from each program about their sources of funding; number of

staff, children, and families served; rates of staff turnover schedule; supervisory structure; and whether the program was already receiving IECMHC services. Because the comparison group consisted of programs operating "business as usual," some programs were receiving a different model of IECMHC.

The need to consider all of this information presented many challenges in trying to understand how to define a "program." Essentially, every program was unique. Programs ranged from independent, privately owned storefront childcare settings to a school district program that served an entire small city and offered occupational, speech, and physical therapy. Some programs blend funding, so including them meant representation of more than one "system." Often programs were large and had multiple classrooms—more than could be served by one consultant given the 10 to 12 hours per month allotted to the program per the Model. All these factors made defining a program and selecting programs a complex and time-consuming process. The variety of programs selected had to meet our goal of using one model across various types of early childhood programs representing different systems. We also wanted to match intervention and comparison programs based on size and participant demographic characteristics to the extent possible.

At the conclusion of the selection process, the sample of programs represented five early childhood systems: Illinois Department of Human Services (IDHS) Center-based Childcare; Head Start; Illinois State Board of Education (ISBE) Preschool for All (PFA); Early Head Start (EHS) home visiting; and ISBE Prevention Initiative (PI) home visiting programs. There were 24 programs: 16 in the intervention group and eight in the comparison group. Because there was a limited number of home visiting programs eligible for the study, the final sample consisted of 18 center-based early childhood programs and six home visiting programs. Again, some of the center-based programs were larger than anticipated and could not be adequately served by one consultant with 10–12 hours. Thus, we either counted a large program as two programs, meaning that it would receive twice as many hours of consultation, or we limited the number of classrooms in the program that would participate in the study. It was also necessary to select two home visiting programs that only had two home visitors rather than our desired minimum of three home visitors.

We matched available programs by program type(s), including funding and general demographics of children served (i.e., if programs served a large proportion of Spanish-speaking families), and randomly assigned them to either the intervention or the comparison group. Although it was not possible to achieve equal representation of each program type across all four communities, we were able to reach the desired number of programs for each program type and for each group assignment (intervention vs. comparison). There were 10 programs in the Peoria region and 14 in the Chicago area; 12 programs were classified as urban and 12 were suburban/rural. The Peoria region had seven programs in the intervention group and three in

the comparison group, while the Chicago region had nine programs in the intervention group and five in the comparison group.

The implementation and evaluation teams made efforts to ensure that the directors and staff of the selected programs had the organizational structure and commitment of program leadership and staff. However, unanticipated turnover of program leaders or staff, funding changes, and new program requirements affected their capacity to participate fully. All but one of the programs, however, completed the 21-month implementation period.

Sample Selection

Program Staff

Once we finalized the selection of programs, we solicited staff lists from the programs and invited current employees who were likely to have contact with a consultant (if the program was in the intervention group) to participate in the evaluation. These individuals included frontline staff (teachers, assistant teachers, and home visitors), supervisors, program directors, and other staff who worked directly with children and families (e.g., paraprofessionals, Family Support Specialists, Family Resource Coordinators). All staff who met these criteria at the time of baseline data collection were part of the sample. In addition, we included any new staff who met these criteria and whose hire date was within 6 months of the start of the evaluation.[11]

A subsample of frontline staff was asked to participate in additional activities. This "focused sample" was selected in consultation with the programs and in consideration of resources for the evaluation. We asked that two staff per program participate in the focused sample. Each center-based program that participated in the evaluation selected two teachers who worked with children ages 3–5 to participate in the focused sample. The specified age range of the children was because the observation tool we used was developed and validated for children ages 3–5. Similarly, each home visiting program selected two home visitors to participate in the focused sample. Participation in all evaluation activities was voluntary.

Children (Center-based Programs)

Prior to baseline data collection, teachers in the focused sample were asked to distribute a consent form to parents asking for their permission to complete child assessments on their children over the course of the study. If the teacher received eight or fewer signed consent forms from parents, the teacher was instructed to complete child assessments for those children.

[11] A power analysis was conducted prior to baseline data collection to estimate the necessary sample sizes to detect effects. To estimate the staff sample size, we used the medium effect size found in the Michigan CCEP evaluation (Egeren et al., 2011) for the matched comparison group study on the Goal Achievement Scale, $\eta_p^2 = .07$. For a cluster analysis, a minimum of 5 providers at each of the 12 center-based programs was sufficient to detect effects.

If the teacher received more than eight signed consent forms from parents, the teacher was instructed to select younger children who were not going to kindergarten next year and children they found to be challenging.[12]

Families (in Home Visiting Programs)

At each data collection time point, each home visitor in the focused sample aimed to recruit two parents to participate in a video-recorded home visit and a brief telephone survey after the visit. Home visitors were given a recruitment script and printed handout to explain the study to the potential parent participants. If a parent agreed to participate, the home visitor connected families with the research team, and a member of the research team provided informed consent to the parents, asking parents to allow a visit to be video recorded and to participate in a survey. Participation was voluntary.

Consultants

The Implementation Team Director, in consultation with other members of the Leadership Team representing state or local agencies that employed consultants, selected 12 mental health consultants with the required credentials to implement the Illinois Model. The consultants were all mental health professionals but with varying backgrounds. Some were independent contractors, while others were employees of mental health or human service agencies. All but one of the consultants were women. As a group, they were racially diverse; six of the consultants were White, three were Black, and two were of other racial/ethnic groups. All but one had/has a master's degree. About two-thirds have degrees in social work or counseling, which the others have their degrees in community mental health, public administration or child development. None of the consultants were familiar with the Illinois model before receiving training as part of the pilot, but about two-thirds had prior experience—on average, 8 or 9 years of experience—as mental health consultants using other approaches.

[12] To estimate the child assessment sample size necessary to detect effects, we used the effect sizes found in the Michigan CCEP evaluation (Egeren et al., 2011) for the pre-post provider-reported Devereux Early Childhood Assessment (DECA) ratings, as the matched comparison group study found no differences. We used a 3-level cluster design for the power analysis. Effect size for DECA Protective Factors was d = .86 (large), resulting in a sample size of 4–6 children per classroom in 12 classrooms. Effect size for DECA Behavior Concerns was d = .57 (medium), resulting in a necessary sample size of more than 60 children per classroom in 12 classrooms. Conners-Burrow et al. (2012) found group differences on the DECA during Year 3 of their evaluation on Behavior Concerns, d = .31, and attachment, d = .26, both of which resulted in needing more than 100 children per classroom assuming these effects in a 3-level cluster design.

Data Collection

Frontline Staff

We administered online surveys to frontline staff (home visitors and teachers) at four time points: baseline, Time 2 (6–8 months post baseline), Time 3 (12–15 months post baseline), and Time 4 (19–21 months post baseline). Staff and supervisors were sent an email with a link to the survey, followed by, at minimum, 5 email reminders. Survey data were collected via REDCap.[13] Table A-1 in Appendix A presents the sample sizes and response rates at each data collection time point.

The surveys contained nonstandardized demographic questions, nonstandardized questions about frontline staff's access to mental health consultation and the quality and frequency of any consultation received, and nonstandardized questions about the supervision they received. Additionally, standardized measures of burnout, depression, reflective capacity, self-efficacy, supervision, and working with children/families with challenging behaviors were also included in the survey. These constructs were important to measure, as these are the intended staff outcomes of the intervention. We selected some measures because they aligned with the theory of change or had been developed for previous IECMHC research, even though they do not yet have published information on their psychometric properties. For example, Shivers (2011) created the Social and Emotional Development Inventory to measure knowledge and skills gained from the "Teaching Pyramid Model" training designed by the Center for Social Emotional Foundations of Early Learning. The standardized measures are briefly described below. (Also see Table A-4 in Appendix A.)

- The *Reflective Functioning Questionnaire* (RFQ; Fonagy et al., 2016) consists of 8 different items which are re-scored to be used in two different subscales, each consisting of 6 items. The initial scale ranges from 1 ("Strongly disagree") to 7 ("Strongly agree"). Each of the 8 items is re-scored into a scale ranging from 0 to 3. The 6 items in each subscale are averaged together so that the Certainty subscale would have a possible score range of 0–3 and the Uncertainty subscale would have a possible score ranging from 0 to 2.33.
- The *Maslach Burnout Inventory* (MBI)-Human Services (Maslach et al., 1996), contains 22 items measuring three facets of burnout in the following subscales: Emotional Exhaustion, Depersonalization, and Personal Accomplishment. The response scale, which is labeled at each point, ranges from 0 ("Never") to 6 ("Every day"). Subscale scores are sums of the item scores, resulting in possible scores ranging from 0 to 54 for Emotional

[13] REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for research studies, providing 1) an intuitive interface for validated data capture; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for data integration and interoperability with external sources (Harris et al., 2019; Harris et al., 2009).

Exhaustion (9 items), 0 to 30 for Depersonalization (5 items), and 0 to 48 for Personal Accomplishment (8 items).

- The *Patient Health Questionnaire* (PHQ; Kroenke et al., 2003) measure of depression consists of two items. The response scale, which is labeled at each point, ranges from 0 ("Not at all") to 3 ("Nearly every day"). The two items are summed, resulting in possible scores ranging from 0 to 6.
- The *Reflective Supervision Rating Scale* (RSRS; Ash, 2010) consists of 17 items. The response scale, which is labeled at each point, ranges from 1 ("Rarely") to 3 ("Almost always"). The 17 items are summed, resulting in possible scores ranging from 17 to 51. The RSRS was not administered to supervisors.
- The *Supervisory Worker Alliance Inventory* (SWAI; Efstation et al., 1990) consists of 19 items for frontline staff and supervisors. These 19 items are used in two subscales: the Rapport subscale consisting of 12 items and the Client Focus subscale consisting of 7 items. The supervisor version asks respondents to complete 7 additional items for the Identification subscale. The response scale ranges from 1 ("Almost never") to 7 ("Almost always"). Each subscale is an average of the items in it, thus possible scores range from 1 to 7 for each subscale.
- The *Goal Achievement Scale* (GAS; Alkon et al., 2003) consists of 13 items (a 14th item was not included in our analysis because it could not be asked at baseline). The response scale, which is labeled at each point, ranged from 0 ("Not at all") to 2 ("Very much"). The 13 items are summed, resulting in possible scores ranging from 0 to 26. With the authors' permission, we adapted the GAS to administer to home visitors.
- The *Teacher Opinion Scale* (TOS; Geller & Lynch, 1999) consists of 12 items. The response scale, which is labeled at each point, ranges from 1 ("Strongly disagree") to 5 ("Strongly agree"). The 12 items are summed together, resulting in possible scores ranging from 12 to 60. With the authors' permission, we adapted the TOS to administer to home visitors. The TOS was not administered to supervisors.

Supervisor and Program Directors

Supervisors were administered surveys at the same four time points with some of the same measures as the staff surveys: nonstandardized demographic questions; nonstandardized questions about their access to mental health consultation and the quality and frequency of any consultation received; and standardized measures of the supervisory relationship (Supervisory Working Alliance Inventory), reflective functioning (Reflective Functioning Questionnaire), depression (Patient Health Questionnaire-2), burnout (Maslach Burnout Inventory), and working with children/families with challenging behaviors (Goal Achievement Scale). Supervisors were

also asked nonstandardized questions about the supervision they provided. Program directors were not administered surveys.

Supervisors and program directors were interviewed by phone at all four time points. Interviews lasted approximately 60–90 minutes and included questions about the program, the strategies supervisors and directors used to support staff in working with children or family members with challenging behaviors and emotions, and their experiences with mental health consultation.

Focused Sample

Teachers

Teachers in the focused sample were asked to participate in interviews, classroom observations, and child assessments, in addition to the surveys all teachers were administered.

Interviews. We conducted interviews with teachers by phone at baseline and Time 3. Each interview was scheduled at a date and time that was convenient for the teacher and lasted approximately 60 minutes. The interview consisted of a series of open-ended questions about the program, teacher strategies used to support social and emotional development, strategies for working with children with a range of behaviors and emotions, and

Box 4. The Climate of Healthy Interactions for Learning & Development (CHILD) Tool

The CHILD (Gilliam & Reyes, 2017) is an observational assessment of the social and emotional (mental health) climate of early care and education settings. A healthy climate is considered one with authentic warmth and a developmentally appropriate and child-centered pedagogy, the equitable treatment of children, and a genuine interest to foster children's psychosocial well-being and holistic development. CHILD items are scored on a five-point scale ranging from -2 to +2. Negative scores indicate a climate that undermines a child's mental health, positive scores indicate one that promotes mental health, and 0 sets the baseline expectation of "doing no harm." The CHILD consists of 28 behavioral items across nine dimensions and two auxiliary dimensions[a] (Reyes & Gilliam, 2018):

1. Transitions: Smooth, efficient, flexible, and productive transitions
2. Directions & Rules: Behavior scaffolding characterized by setting, modeling, and enforcing clear, consistent, and developmentally appropriate rules and applying positive behavior strategies
3. Social & Emotional Learning: Fostering in children self-awareness, self-management, relationship skill-building, social problem-solving, and responsible decision making
4. Staff Awareness: Staff monitoring and attunement to both overt and subtle signals from children
5. Staff Affect: Emotional state and well-being of staff
6. Staff Cooperation: Staff demonstration of teamwork and genuine enjoyment of each other
7. Staff–Child Interactions: Staff interactions with children characterized by dignity, respect, genuine relationships, equity, and the celebration of diversity
8. Individualized & Developmentally Appropriate Pedagogy: Staff fosters holistic development through a child-centered and individualized approach
9. Child Behaviors: Facial expressions, body language, and interactions with peers and adults
10. Equity: Staff are aware of all children and attend to individual needs (auxiliary dimension)
11. Warmth: Staff honestly enjoy their work and time with the children and the children like being with the staff (auxiliary dimension)

[a] Auxiliary dimensions include items from the above dimensions and are not unique to either Equity or Warmth.

experiences with mental health consultation. The interviews also included the Provider Reflective Process Assessment Scales (PRPAS; Heller, 2017). The PRPAS consists of 14 items in six scales that correspond to different dimensions of reflective capacity: Self-Knowledge, Self-Regulation, Multiple Perspectives, Collaboration, Process, and Authentic Attitude. (See Table A-5 for description of the PRPAS.) The response scale ranges from 0 ("Reverse") to 4 ("High"). The items in each scale are summed, resulting in possible scores ranging from 0 to 12. Interviews were audio recorded and transcribed for analysis. Incentives were provided for participation.

Classroom observations. Observations were conducted in the classrooms of the teachers in the focused sample at baseline, Time 2, Time 3, and Time 4. A member of the research team worked with the program director and the teacher to identify the best date and time for an observation of a 2-hour block of a "typical" day. The observation tool used was the Climate of Healthy Interactions for Learning & Development (CHILD; Gilliam & Reyes, 2017), as described in Box 4. The research team was trained on the tool by one of the measure's developers, who also provided refresher trainings throughout the study. The CHILD is scored across four observation blocks of 20 minutes each, with a 10-minute break after each to score the block. Twenty percent of the classrooms were observed and coded by two members of the research team for reliability. Researchers entered both original scores to calculate reliability, and the observers discussed any discrepancies to agree on a consensus score, which was used for analysis.

Child assessments. Up to eight children in each of the classrooms of the focused sample of teachers were selected for the child assessments sample. The child assessments were completed by teachers via online survey and included three standardized scales to measure social-emotional development and behavioral dysregulation: Strengths & Difficulties Questionnaire (SDQ; Goodman et al., 1997); the Preschool Expulsion Risk Measure (PERM; Gilliam & Reyes, 2018); and the Devereux Early Childhood Assessment (DECA-P2; LeBuffe & Naglieri, 2012) version for ages 36–60 months. We used a brief modified version of the SDQ containing six items (Perry, 2013; Stephan et al., 2011). Only respondents who indicated that a child has difficulties with emotions, concentration, behavior, or getting along with other people were administered the remaining items to assess the severity of the problems. Three items asking if these difficulties upset the child, if these difficulties interfere with the child's peer relationships, and if these difficulties interfere with the child's learning are summed, resulting in an Impact Score, which can range from 0 to 6.

The PERM (Gilliam & Reyes, 2018) measures teacher perception of disruptive child behavior in four subscales: classroom disruption (the degree to which a child's behaviors create disruptions in the classroom); fear of accountability (the degree to which children's behaviors may pose a risk of injury for which the teacher might be accountable); hopelessness (the degree to which the teacher may feel hopeless that anything can be done to improve behaviors in the

classroom); and teacher stress (the degree to which children's behaviors are associated with increased teacher stress). The PERM also provides a total score.

The DECA-P2 (LeBuffe & Naglieri, 2012) is a standardized behavior rating scale that provides an assessment of behavioral concerns and within-child protective factors related to resilience. The DECA-P2 includes three scales assessing protective factors: Initiative, Self-Regulation, and Attachment/Relationships, as well as a Total Protective Factors scale. It also includes a Behavioral Concerns scale.

Home visitors. The focused sample of home visitors was asked to participate in interviews, video recordings of home visits, and a brief survey about the video-recorded home visit.

Interviews. The interviews with home visitors were conducted by phone at baseline and Time 3. They lasted approximately 60 minutes. The interview consisted of a series of open-ended questions about the program, home visitor strategies used to support social and emotional development in children and families, and strategies for working with challenging behaviors and emotions. The interviews also included the PRPAS (Heller, 2017), explained in more detail earlier. Interviews were audio recorded and transcribed for analysis. Incentives were provided for participation.

Home visit observations. Home visitors in the focused sample were asked to video record two home visits, each with a different family, at all four observation time points. The research team used the Home Visit Rating Scales-Adapted & Extended to Excellence (HOVRS-A+; Roggman et al., 2010) to assess the home visitor's practice and the relationship between the home visitor and parent. This instrument has been validated in various program models with

Box 5. The Home Visit Rating Scales-Adapted & Extended to Excellence (HOVRS-A+)

The HOVRS-A+ (Roggman et al., 2010) is a widely used observation tool to assess home visitors' strategies and relationships during home visits, such as the home visitor's effectiveness in engaging the caregiver and child in the visit. It consists of the following seven scales, with higher ratings reflecting more responsive behaviors on the part of the home visitor:

1. *Home Visitor Responsiveness to Family:* Assesses preparedness for home visit, responsiveness to parent and child, and efforts to get information and input from parent
2. *Home Visitor-Family Relationship:* Examines relationship between the home visitor and the family, including warmth, positive interactions, and respect for family
3. *Home Visitor Facilitation of Parent-Child Interaction:* Assesses home visitor at facilitating and promoting positive parent–child interactions during the home visit.
4. *Home Visitor NonIntrusiveness/Collaboration with Family:* Focuses on the lack of intrusiveness by the home visitor on parent behavior and parent–child interactions during the visit.
5. *Parent–Child Interaction during Home Visit:* Examines the nature of the parent-child relationship, as observed during the home visit.
6. *Parent Engagement during Home Visit:* Examines the engagement of the parent and the activities of the home visit.
7. *Child Engagement during Home Visit:* Focuses on the child's engagement—involvement and interest—in the activities of the home visit.

different cultures, such as Spanish-speaking Latino families, rural White families, and urban Black families (Schodt et al., 2015). See Box 5 for a description of its scales. Two pairs of coders (one pair to code English home visits and one pair for visits conducted in Spanish), all blind to the condition of the home visitors, established inter-rater reliability with a Kappa of 0.98 and met to discuss their ratings and any discrepancies throughout the process.

Post video-recorded home visit survey. After video recording a home visit, we sent the home visitor an online survey to complete to provide additional information about the visit. Questions focused on who was present, the goals of the visit, characteristics of the child and parent, length of time in the program, and whether it was a typical visit.

Parent survey. Parents participating in the home visiting video recordings were also administered a survey. The survey contained demographic questions and standardized measures of parenting and depression. Five subscales from the Healthy Families Parenting Inventory (HFPI; LeCroy & Milligan Associates, 2004) were included in the parent survey: Parent/Child Interaction, Home Environment, Role Satisfaction, Parenting Efficacy, and Problem Solving. The HFPI was specifically developed to measure families in home visiting programs, and each subscale can be administered on its own (Krysik & LeCroy, 2012). In addition, a brief, validated measure of depressive symptoms, the Patient Health Questionnaire-2 (PHQ-2; Kroenke et al., 2003) was included in the parent survey. For families with infants under 12 months of age, an additional tool was administered in the survey to measure their perception of infant crying, the Infant Crying & Parent Well-Being Screening Tool (Katch & Burkhardt, 2021). Parents completed the survey either over the phone or online via REDCap.

Consultants

Consultants were an important source of information to assess the quality and fidelity of implementation. The consultants used an online database called the Consultant Log, starting with their first contact with a program and continuing throughout the course of the pilot until they transitioned to the intermittent support period. They recorded their hours, activities, the individuals with whom they met, the topic of the consultation, success, challenges, and next steps to follow-up on the activities. The implementation team provided Chapin Hall researchers copies of these consultant logs for analysis. Each log entry contained the date and number of hours spent in consultation, the type of consultation (e.g., reflective consultation, observation, team meeting), the content and issues raised during the consultation, and the successes, concerns and next steps identified for the program or case.[14]

[14] Previous evaluations of IECMHC have used consultant logs to varying degrees of success; some researchers have reported low response rates and late implementation of the logs, preventing them from being a full and accurate representation of implementation (Egeren et al., 2011).

The final set of consultant logs for the 15-month intensive implementation period included 954 entries, 765 (80%) for the early childhood center-based programs and 189 (20%) for the home visiting programs. The content or issues recorded in the logs varied in terms of quality of information. Consultants sometimes provided detail about the topic of a consultation, such as a conflict with a coworker or a challenge communicating with a family. Other times, they only reiterated the type of consultation—reflective consultation/supervision or observation, for instance—without further detail. In most cases, coders were able to use other information in the logs to discern the level and focus of the consultation, but they were unable to categorize 101 entries for early childhood center-based programs and 22 for home visiting programs.

Survey Sample Characteristics at Baseline

In Spring 2018, we sent baseline surveys to 293 staff; 195 (67%) responded.[15] Of those 195 teachers and home visitors, 136 (70%) completed at least one other survey and were included in the analyses of change over time. We refer to this group of 136 as the analytic sample. Of those 136, 72 were in the intervention group and 64 were in the comparison group. Table 2 presents their demographic characteristics as well as those of the analytic sample of supervisors followed over time. The sample sizes and response rates for the surveys and other data sources at all data collection time points can be found in Table A-1 in Appendix A. Table A-2 and Table A-3 show how the analytic sample compares to the full baseline sample. There were no statistically significant differences between the full sample and the analytic sample at baseline.

Demographic Characteristics

Frontline Staff

The majority of the staff in the analytic sample identified as female. Staff in the intervention and comparison groups were quite similar in terms of race and ethnicity. About half of each group of staff identified as white, about one-quarter identified as Latino/Hispanic, and about one-quarter identified as Black. In the analytic sample, the intervention and comparison groups were similar in age. Staff in the intervention group had somewhat higher educational levels, however, compared to staff in the comparison group. A chi-square test for staff education in the analytic sample was significant (χ^2=11.84, p = .008). Staff in the intervention group had more participants who had bachelor's degrees, while the education level for the staff in the comparison group was more evenly distributed among the four categories (some college, associate's degree, bachelor's degree, master's degree). On average, staff in both groups have 10 or more years of experience working with children and families while staff in the comparison group have slightly more experience than staff in the intervention group.

[15] Of the 195 staff, 180 completed the full survey, while 15 completed only part of the survey.

Finally, since most of the programs in the study were early childhood center-based programs, the majority of staff in both the intervention and comparison groups worked in center-based early childhood programs. Just a small percentage were home visitors. The sample was evenly distributed in terms of location. Approximately half of staff in intervention programs and comparison programs worked in the Chicago area, while the other half worked in the Peoria area.

Supervisors

Compared to the staff sample, the sample of supervisors for this study was quite small. In Spring 2018, we sent baseline surveys to 31 supervisors; 26 (87%) completed the survey. Of those 26 supervisors, 19 completed at least one additional survey and were included in the subsequent analyses of outcomes. Of those 19, 14 were in the intervention group, and the remaining five were in the comparison group.

Most of the supervisors in the analytic sample identified themselves as female. As a group, the supervisors were diverse in race and ethnicity and age. A majority of the supervisors in both the intervention and comparison groups had earned their master's degree or higher. Supervisors varied widely in terms of their years of experience working with children and families, but, on average, they had at least 12 years of experience.

Finally, like the staff sample, most of the supervisors in the intervention and comparison groups worked in ECE settings, while a small percentage were in home visiting programs. Although there were no significant differences between the two groups, the supervisors in the intervention and comparison groups differed somewhat with respect to location. More supervisors in the comparison group were located in Chicago than in Peoria at baseline.

Table 2. Demographic Characteristics of Analytic Samples at Baseline[a,b]

Characteristic	Staff		Supervisors	
	Intervention *N* = 72	Comparison *N* = 64	Intervention *N* = 14	Comparison *N* = 5
Gender (%)				
Female	96	100	86	100
Male	4	0	14	0
Race/Ethnicity (%)				
Black	20	22	17	40
White	54	48	50	40
Latino(a)/Hispanic	26	30	25	20
Other	0	0	8	0
Age (%)				
Under 20 years	0	5	0	0
20–29 years	24	17	0	0
30–39 years	30	30	29	40
40–49 years	28	17	57	20
50 or older	18	31	14	40
Education (%)				
Some college/no degree	14	24	0	0
Associate's degree	20	32	14	20
Bachelor's degree	57	27	21	20
Master's degree or above	9	16	65	60
Years of Experience[c]				
Mean (*SD*)	10.4 *(9.66)*	11.2 *(8.21)*	13.4 *(6.27)*	18.5 *(10.08)*
Range	1–43	1–39	5–26	8–32
Program Type				
Early Childhood Center	83	94	79	80
Home Visiting	17	6	21	20
Location				
Chicago area	49	53	57	60
Peoria area	51	47	43	40

[a] Percentages reflect those who responded to the survey question.
[b] Analytic sample refers to the sample of staff who responded at baseline and at least one other time point and, therefore, could be included in analyses of change over time.
[c] Sample sizes for Years of Experience were as follows: $n = 64$; $n = 46$; $n = 11$; $n = 4$

Data Analysis

Quantitative

The data analysis plan was to test each outcome (each scale from the surveys and observations) as the dependent variable and the group assignment (intervention or comparison) as the primary independent variable, along with covariates. Thus, we planned to conduct 54 analyses to test the effect of the intervention. Because the data were clustered and longitudinal, we

analyzed the staff surveys, classroom observations, child assessments, and home visiting observations using linear mixed modeling (LMM).[16] LMM is a parametric linear model used to analyze clustered, longitudinal, or repeated-measures data, including both fixed-effect parameters and random effects factors. In particular, the fixed-effect parameters describe the relationships of the variables included in the analysis to the dependent variable for an entire population and the random effects are specific to the clusters identified in this analysis (West et al., 2007). Additional details about the LMM approach and the variables included in the final models can be found in Appendix A.

The staff survey data, classroom observations, and child assessments are clustered at the program level and include repetitive measures post-baseline. This methodology proposed by West et al. (2007) involves the following steps: 1) fit a model with a "loaded" mean structure, 2) select a structure for the random effects, 3) select a covariance structure for the residuals, and 4) reduce the model by removing covariates, if needed. Thus, we initially included all possible covariates—that is, all program-level and staff-level variables—for each outcome. Then we removed any variables that prevented the model from running because they were not statistically significant or were highly correlated with other variables in the model.

We administered the survey to staff at baseline, Time 2, Time 3, and Time 4; however, not all staff had complete survey data. The surveys with demographic and baseline data and at least one other point in time (Time 2, Time 3, or Time 4) were included in the final dataset. In addition, because there are different numbers of staff per program, we have an unbalanced clustered dataset. LMM can be fitted to this type of data structure with the assumption that any missing data are missing at random. For the staff survey data, we estimated an LMM with clustered longitudinal data with three levels:

- Level 3 - Cluster of Clusters: Programs
  There are 19 programs in this analysis.
- Level 2 - Unit of Analysis: Staff
  There are 136 teachers and home visitors in our sample.
- Level 1 - Time: Measures at Four Points in Time (including baseline)
  Each of the dependent variables included in this analysis was measured at four points in time.

For the classroom observation data, we estimated an LMM with clustered longitudinal data with three levels:

- Level 3 - Cluster of Clusters: Programs
  There are 11 programs that are part of this analysis.

[16] LMM is also referred to as hierarchical linear modeling (HLM) or multilevel modeling (MLM). We could not analyze the supervisor survey data with LMM because the sample was too small.

- Level 2 - Unit of Analysis: Classrooms/Teachers
  There are 28 classrooms/teachers in our sample.
- Level 1 - Time: Measures at Four Points in Time (including baseline)
  Each of the dependent variables included in this analysis was measured at four points in time.

For the child assessment data, clusters occurred at the classroom level, and classrooms were clustered at the program level. The child assessment data included repetitive measures at two points in time post-baseline. Given the structure of the child assessment data, we estimated an LMM clustered longitudinal data with four levels:

- Level 4 -Cluster of Clusters: Programs
  There are 12 programs that are part of this analysis.
- Level 3 - Cluster of Clusters: Classrooms
  There are 21 classrooms/teachers that are part of this analysis.
- Level 2 - Unit of Analysis: Students
  There are 136 students in our sample.
- Level 1 - Time: Measures at Three Points in Time (including baseline)
  Each dependent variable included in this analysis was measured at three time points.

The home visiting observations contained different families at each observation, thus the data did not contain repetitive measures at the family level. We analyzed the three-level hierarchical linear models for each of the dependent variables in the home visiting observation data. This methodology involves the following steps: 1) fit the initial unconditional (variance components) model and decide whether to omit the random home visitor effects, 2) build the level 1 model by adding parent-level covariates, 3) build the level 2 model by adding home visitor covariates, and 4) build the level 3 model by adding program-level covariates. There are different numbers of home visitors per program as well as different numbers of families who work with a specific home visitor. Thus, we have an unbalanced clustered dataset. LMM can be fitted to this type of data structure with the assumption that any missing data are missing at random. For the home visiting observation data, we estimated an LMM with clustered longitudinal data with three levels:

- Level 3 - Cluster of Clusters: Programs
  There are 5 programs that are part of this analysis.
- Level 2 - Cluster of Units: Home Visitors
  There are 7 home visitors in our sample.
- Level 1 - Unit of Analysis: Families
  There are 42 families in our sample.

Because home visitors recorded visits with different families at each time, different parents completed the parent survey at each time. The parent survey data were not repeated measures

but they were clustered hierarchically. Thus, we conducted hierarchical two-way ANOVAs, with the program nested within group over time, to compare any differences in the scores of the parents whose home visitors were receiving the intervention compared the parents whose home visitors were not receiving the intervention. We used a multilevel analysis to account for the fact that some families had home visitors from the same program.

- Level 2 - Cluster of Clusters: Programs
  There are 5 programs that are part of this analysis.
- Level 1 - Unit of Analysis: Families
  There are 51 families in our sample.

The supervisors' surveys were analyzed separately due to the small sample size (N = 19 with 14 at intervention programs and the remaining 5 at comparison programs) and the absence of other quantitative data to include.

Interviews

Interview transcripts were imported into Atlas.ti software for analysis. We analyzed the transcripts thematically using open and descriptive coding (Saldaña, 2015). Two members of the qualitative research team read a sample of the same transcripts for each respondent group (home visitors, teachers, supervisors, and directors) to develop initial codebooks for each respondent group.[17] The researchers met regularly to discuss coding, create additional codes as needed, and resolve discrepancies. Once the interviews were coded, the coded material was exported to Excel to discern themes within and across respondent groups. Our analyses were largely guided by our research questions, though we allowed new topics or themes to emerge when relevant. Regular research team discussions were held to assess the interpretation of narratives (i.e., construct validity). We established the validity of themes and key findings by triangulating our data sources (i.e., surveys, follow-up interviews, supervisors, home visitors, and consultant interviews).

Consultant Logs

We used the data from the consultant logs to analyze the extent to which the Illinois Model was implemented as intended. Our analysis of implementation considered both structural and process indicators. Structural indicators include dosage, or the number of hours of consultation program staff received, and adherence, or the alignment between the types of consultation provided with the activities recommended by the model. We also looked at the topics or issues

[17] Consultant interviews were subject to a similar but different process. The codebooks for the other respondent groups served as a guide for the consultant codebook. Two members of the qualitative research team reviewed and coded the same transcript for reliability.

raised in consultation. In brief, we analyzed the logs in terms of the following three aspects of implementation:

- Dosage of, or exposure, to consultation: What was the total number of hours of consultation provided to each program?
- Alignment of activities to the model: What types of activities and with whom did the consultant engage?
- Content of activity or issues raised in consultation: Was the content of consultation in line with expectations to be program-, classroom-, and case-focused?

Dosage

In consultation with the Evaluation Team, we determined that consultants who were able to complete at least 80% of their expected hours would have satisfied the expected number of hours established for this pilot study. For typically sized early childhood programs and the one large home visiting program, the expectation was 10–12 hours per month, and we used 10 hours as the standard for calculating completed hours.[18] For small home visiting programs, for which the expectation was 5–7 hours per month, we used 5 hours per month.

Because implementation was staggered, start dates for the programs differed: Group 1 started consultation in April 2018; Group 2, in August 2018; and Group 3, in November 2018. Thus, we used each program's start date to calculate the percentage of goal hours met by each program. We excluded the summer months in calculating the percentage of goal hours met for programs that closed during the summer.[19]

Alignment

In addition to dosage, it was also important to understand whether the consultants' activities were in line with the recommended practices identified by the Illinois Model developers. Preparing the consultant log data for analysis of consultant activities took considerable time because of variations in the quality and amount of information in the logs. Not all consultants reported their hours and activities in the same way. To ensure the data were as complete and accurate as possible, we communicated with individual consultants to fill in missing information on their hours and types of activities to categorize activity descriptions and hours as consistently as possible.

Content

In the analysis of the content of the consultant logs, we used the definitions developed by the implementation team for the three types of consultation—programmatic consultation,

[18] For two larger programs that are counted as two programs for the evaluation, we used 20 hours/month.
[19] Six programs are fully closed in the summer, 3 have lighter summer schedules, and 14 are year-round programs.

classroom and home consultation, and child and family consultation. (See Box 3 on page 14.) The analysis involved three steps: (1) cleaning the data (e.g., removal of duplicate information and resolving missing data); (2) identifying the level of the issues; and (3) categorizing the issues and identifying emerging themes. The amount and quality of information in the consultant logs varied across the sample of consultants, but we were able to assign the vast majority of the entries to one category.[20]

Two researchers initially selected a set of sample excerpts from the logs and confirmed the applicability of these three themes to the log entries for all types of consultation activities (e.g., team meeting, reflective consultation, or observation) and for both types of programs (i.e., center-based or home-based). One researcher assigned one of the three levels to each log entry; a second researcher checked the assignment for consistency and accuracy. Conflicts were resolved by a third researcher or through team discussion and consensus.

If the only information provided was the form of the activity (e.g., observation or family event), some of the categories were decided based on the activity rather than the content of the issues raised. For example, the research team assigned the classroom/home level to general classroom observations conducted by the consultant, and the child/family level to observations that addressed issues regarding specific children or families. Similarly, we assigned family events or staff training that the consultant facilitated to the program level. To further identify the focus or theme of the consultation, the researchers revisited the logs and discussed potential issue categories. One researcher reviewed a set of samples from the consultation logs and identified potential categories; a second researcher checked the categories and combined or narrowed down multiple categories as necessary.

Once primary categories were developed, one researcher assigned a category based on the main issues reported in the log (Category 1). Multiple categories were assigned (up to Category 4) when applicable. A second researcher checked the assignment for accuracy. Next, the researchers combined similar categories (e.g., "communication with parents" and "family engagement") and broke down a category that included multiple topics (e.g., "administrative" and "staff turnover"). The researchers created several categories that addressed forms or activities of the consultation (e.g., observation, training) rather than the content of the issues, depending on the information provided in the log. For the categories that were linked to only one or two log entries, the researchers agreed on whether to combine the categories or keep them as independent categories as long as these "miscellaneous" categories informed the focus or theme of the issues discussed at each level.

[20] Of the 954 entries, 68 (6%) could not be coded because of insufficient information. Seventeen entries for center-based programs were double-coded because they reflected more than one level (e.g., program and child) but we were to assign all others to a single category.

When the log contained a limited amount of information regarding the content of the issues raised, researchers considered additional information in the logs regarding the consultation—such as how the consultant responded to the needs of the staff/program and what competencies or plans of action were addressed—to determine the level and category/categories of the issues.

Chapter Summary

The evaluation of the Illinois Model of IECMHC required a rigorous design that was also flexible and responsive to the community and program characteristics of the sample. Like most approaches to IECMHC, the form and content of consultation in the Illinois Model depends on the needs and goals of the program staff, the relationship between the consultant and staff, and a number of other factors, including organizational context and how early childhood staff understand and use the consultants' support. This made studying implementation—and the fidelity of implementation—complicated. Although the model is based on prior research and tools developed by the field, the Illinois Model is unique—particularly in the extent to which it emphasizes the development of reflective capacity in staff and supervisors—and required a unique design and a mix of quantitative and qualitative methods.

Findings: Implementation of the Illinois Model

> Before we started [consultation] we would have reflective supervision on the calendar, then something would come up, and we'd reschedule. And then it'd be two months down the line and we still hadn't rescheduled [supervision]. We didn't know what we were doing, so it was just something that we would do, but we weren't really doing it. Then once we had [the consultant] come on, it's been like, "Oh, we have this time, we're using this time, we're not changing the schedule. So it's been really good. And it's nice to have [the consultant] model all of those reflective pieces. —*Program director*

Understanding the implementation fidelity and quality of the Illinois Model is critical to understanding its success in achieving its intended outcomes (e.g., Daro, 2010; Durlak, 2015; Fixen, et al. 2009; Hansen, 2014). However, it is challenging to assess the fidelity of a multidimensional, flexible intervention like the Illinois Model because its effectiveness depends on program characteristics, the responsiveness of staff, and consultants' ability to develop relationships with staff and provide the services expected by the model.

We analyzed both structural and process indicators of fidelity to study implementation. Primary structural indicators were *dosage*, or the number of hours of consultation delivered to a program, and *adherence*, or the alignment between the types of consultation provided with the activities recommended by the model. Data for these indicators came primarily from the consultant logs, an online database of consultant activities, recorded throughout the 15-month intensive implementation period. We considered programs that received at least 80% of their expected hours to have received the minimum number of hours required by the model. For typically sized early childhood center-based programs and a large home visiting program, the expectation was 10–12 hours per month, and we used 10 hours as the standard for calculating completed hours.[21] For small home visiting sites, for which the expectation was 5–7 hours per month, we used 5 hours per month.

We supplemented the information in the consultant logs with data from surveys and interviews. Qualitative interviews were particularly helpful in describing the quality of consultants' relationships with staff and supervisors and identifying some of the many factors that may have

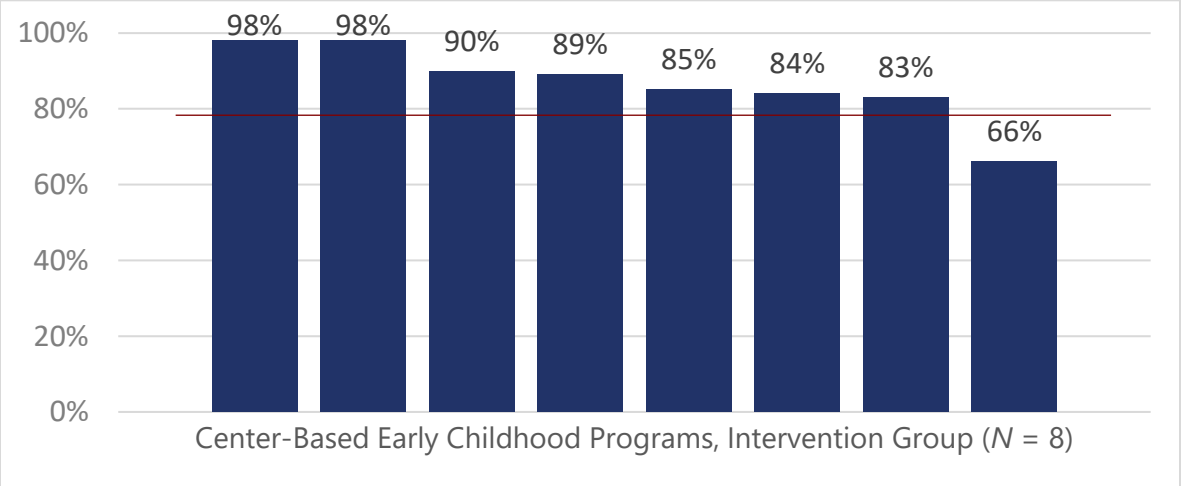
[21] For two larger sites that are counted as two sites for the evaluation, we used 20 hours/month. We excluded the summer months in calculating the percentage of goal hours met for programs closed in the summer.

affected the variation in consultation hours and activities, such as program size and structure, funding changes, staff readiness to work with the consultant, and staff turnover. In the following sections, we present our findings regarding implementation, with a focus on fidelity, the content of consultation, then implementation barriers and facilitation.

Dosage: Hours and Frequency of Consultation

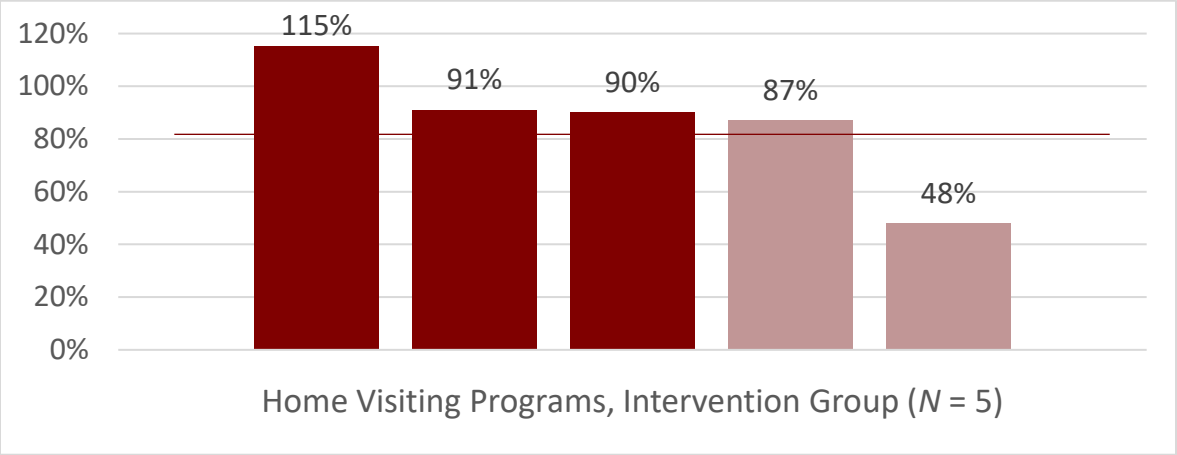
As seen in Figure 3 and Figure 4, all but two of the Intervention sites received greater than 80% of their expected goal hours, adjusted for program size and summer schedules. One was a school-based early childhood center-based program, where it was very difficult to make time in the schedules of the program director and teachers. The other was a small home visiting program that experienced a complete staff turnover soon after the start of the pilot.

Figure 3. Percent of Expected Consultation Hours Received by Eight Early Childhood Center-based Programs (15-month Intensive Period)*



*Full implementation was defined as meeting 80% or more of the hours expected, adjusted for program size and summer schedules. The expectation for mid-sized programs was 10 hours/month.

Figure 4. Percent of Expected Consultation Hours Received by Five Home Visiting (15-month Intensive Period)*



*Full implementation was defined as meeting 80% or more of the hours expected, adjusted for program size and summer schedules. The expectation for mid-sized programs was 10 hours/month.

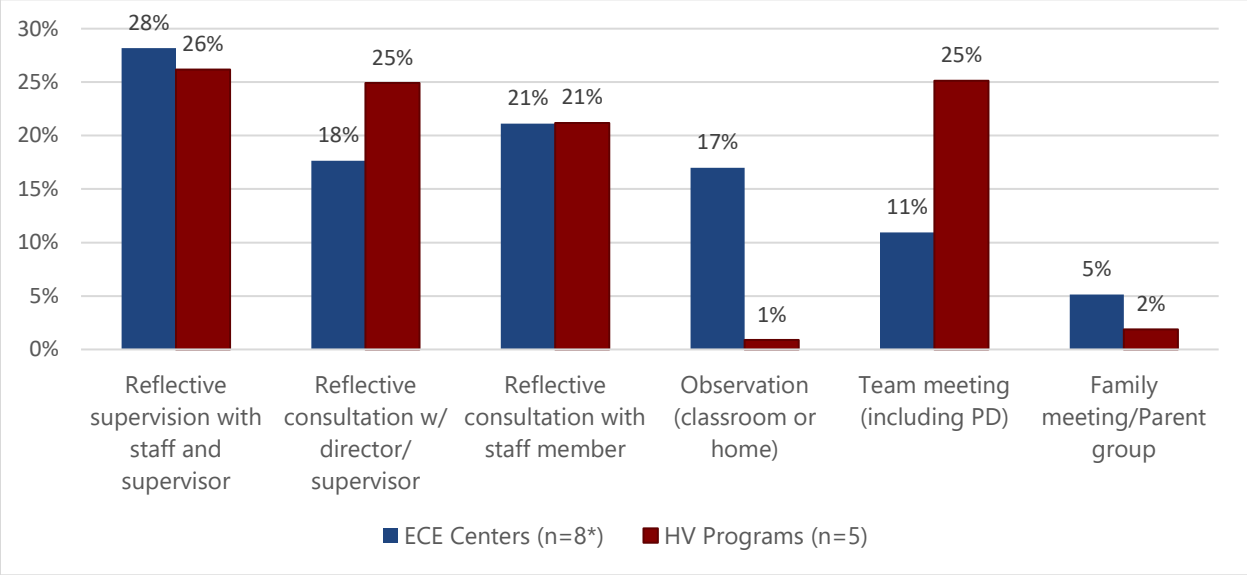
Adherence: Types of Consultation Activities

Overall, the specific consultant activities adhered to the expectations of the Illinois Model, as shown in Figure 5. All the intervention programs received the expected types of consultant services, although the distribution of activities varied among the programs. In accordance with the model, the majority of activities in both early childhood center-based and home visiting programs were reflective consultation activities—reflective consultation with supervisors and staff together, reflective consultation with supervisors and directors, and reflective consultation with individual staff. Consultants in home visiting programs recorded a somewhat higher percentage of reflective consultation sessions with directors and supervisors than early childhood center-based programs. This finding was not surprising given the challenges reported in qualitative interviews regarding scheduling time with supervisors in sometimes understaffed center-based programs.

Both consultants and supervisors reported challenges arranging meetings with supervisors and individual staff together—a highly recommended practice of the Illinois Model—but early childhood center-based programs and home visiting programs were able to arrange these meetings with equal frequency. It was much more common for consultants at early childhood center-based programs to spend time observing in classrooms than for them to observe a home visit. Consultants conducted observations to help them understand the classroom context and needs of teachers as well as to observe individual children. Home visiting programs were more likely to hold team meetings on a regular schedule than were early childhood center-based programs, which is one explanation for the differences between the two types of programs in the proportion of time consultants spent in group meetings. Finally, as shown in Figure 5, consultants spent little time meeting with parents. The Illinois Model recommends that

consultants meet with parents only with another staff member and, more often than not, consultants followed this practice. There were a few occasions, however, when directors or supervisors requested that consultants meet with parents individually.

Figure 5. Percentage of Consultant Activities by Program Type (15-month Intensive Period)



Content and Process of Consultation

Table 3 and Table 4 describe the themes or categories of the main issues raised during mental health consultation sessions, as recorded by the mental health consultants in their work with staff and supervisors during the intensive implementation period. The majority of the issues or content of consultation activities at early childhood center-based programs were at the program or classroom level, while most of the activities at home visiting programs were at the program level. Although we provided definitions of the three levels previously, we repeat them here to make interpretations of the findings easier.

- *Programmatic consultation*: Assesses a program's structures, policies, procedures, professional development opportunities, philosophy, mission, and practices as they relate to supporting the mental health of young children and their families.
- *Classroom and home consultation*: Collaborates with parents and staff to assess relationships, routines, and practices that impact the classroom or home climate.
- *Child and family consultation*: Collaborates with families, staff, and other caregivers to understand and respond effectively to a child's mental health needs. Assists caregivers and home visitors to understand and effectively respond to the mental health needs of a family. Consults with families, staff, and other caregivers about a particular child or family.

Table 3. Content of Consultation in Consultant Logs for Nine Early Childhood Center-based Programs[a]

Category of content/issues raised[b]	Program level (*n*)	Classroom/Home level (*n*)	Child/Family level (*n*)	Total	
				n	%
Reflective supervision/ consultation with staff	64	83	39	186	29.6%
Observation	0	119	32	151	24.0%
Reflective supervision/ consultation with supervisor	45	1	5	51	8.1%
Staff dynamics (communication, relationships, conflicts, collaboration)	45	2	0	47	7.5%
Administrative/program management (e.g., internal program issues, staff turnover)	35	4	0	39	6.2%
Child behavior	0	3	26	29	4.6%
Family support	4	0	25	29	4.6%
Introduction to MHCI/Pilot	24	0	0	24	3.8%
Training	22	1	0	23	3.7%
Staffing and staff performance	14	0	0	14	2.2%
Parent engagement/ communication/family event	13	0	1	14	2.2%
Trauma (e.g., gun violence and other community violence)	0	1	11	12	1.9%
Personal/life stressors (e.g., health, housing, marital issues)	0	10	0	10	1.6%
Grand total[c]	**266 (42%)**	**224 (37%)**	**139 (21%)**	**629**	

[a] Table includes entries for one program that did not complete the pilot.
[b] Another 58 log entries were categorized as "no information available" and 43 logs as "unclear" because of limited information.
[c] Seventeen entries were double-coded as both program and classroom level.

Consultants also recorded a number of issues related to program management and other administrative matters. This category includes policies or procedures in need of clarification, for example, policies for working with sick children, filing systems, gathering information about new/perspective families, etc. Other issues in this category include planning for orientations, staff role clarification, leadership changes, and the impact of program growth on its day-to-day operations. The consultants in the home visiting programs spent more time focused on program-level issues, particularly staff turnover and recruitment and time management. Meanwhile, the consultants in early childhood center-based programs spent more time focused on child behavior and family support than their counterparts in home visiting programs.

Table 4. Content of Consultation in Consultant Logs for Five Home Visiting Programs

Category of Content/ Issues Raised[1]	Program level	Classroom/ Home level	Child/Family level	Total[2]	
				n	%
Reflective supervision/ consultation with Staff	14	8	30	52	31.5%
Administrative/program management (e.g., internal program issue, staff turnover, recruitment of families)	40	1	2	43	26.1%
Introduction to MHC/Pilot	12	0	0	12	7.3%
Explaining value of Reflective Supervision	10	0	0	10	6.1%
Reflective supervision/ consultation with Supervisor	10	0	0	10	6.1%
Professional development/ Training	9	0	0	9	5.5%
Staff dynamics (relationships, conflicts)	9	0	0	9	5.5%
Staffing and staff performance	8	0	0	8	4.8%
Observation	6	0	0	6	3.6%
Trauma (e.g., gun violence and other community violence)	3	0	0	3	1.8%
Family event/parent education (e.g., self-care, child discipline, nutrition)	2	0	0	2	1.2%
Personal care/life stressors	1	0	0	1	0.6%
Grand Total	**122 (74%)**	**11 (7%)**	**32 (19%)**	**165**	

[a] Another 17 log entries were categorized as *no information available* and 5 logs as *unclear* because of limited information.

Table B-1 presents examples of three most frequent categories of issues that consultants in the early childhood center-based and home visiting programs recorded in their logs. Although there were differences between the two types of programs, the most frequent issues were similar. These included: (1) issues related to program development and program management; (2) reflective supervision/consultation with staff; and (3) reflective supervision/consultation with supervisors. In both types of programs, consultants most often recorded reflective supervision/consultation with staff as the purpose or content of the interaction. They also recorded reflective supervision/consultation with supervisors frequently, especially at the early childhood center-based programs. Again, some of these interactions followed the Illinois Model recommendations and involved staff and supervisors together; in other cases, the consultant met separately with supervisors and staff.

Although the consultant logs provided a useful overview of the issues addressed by consultants in their work, interviews conducted with consultants, staff, supervisors, and directors offered more details about the content and context of the consultation. Though there were some clear commonalities in the focus of consultation across the programs in the intervention group, the interviews highlighted the varying needs of the programs and how their needs informed the consultants' work.

Three major consultation content themes emerged in the analysis of the interview data, which align with the themes in consultant log entries: 1) reflective practice, 2) working with children and families, and 3) work relationships. Other themes observed less frequently in the interviews included: training and professional development; personal issues; diversity, equity, and inclusion (DEI) issues; grant compliance; and connections to resources.[22] In the following sections, we discuss these themes and representative excerpts from the interviews. It should be noted that although we discuss these themes separately, they overlapped considerably; for example, building staff's reflective capacity also strengthened their relations with families and children and with coworkers.

Reflective Practice

All of the intervention programs broadly addressed "reflective practice" during their mental health consultation. This theme includes building the capacity of staff to think critically, consider multiple perspectives, and solve problems on their own; helping staff self-regulate; processing home visits with a reflective lens; and processing school/classroom issues with a reflective lens. A major part of the consultants' work was to increase staff understanding, knowledge, and skills through building reflective capacity and practice. Although reflective supervision is considered a "best" practice in home visiting and often woven into home visitors' practices, familiarity and comfort levels with it vary. The same holds true at early childhood center-based programs, with most programs not having a reflective supervision component as part of their model at all. In both individual and group meetings with supervisors and staff, consultants reported discussing the meaning of reflective practice, its importance, and staff comfort level and experience with it.

[22] Note that during our analyses we started to quantify how many sites referenced particular themes and topics as they discussed the content of their consultation. However, we concluded that quantifying in this way was misleading due to our not having interviewed all staff members at each program, attrition among staff members, and the fact that not having a specific quote about a specific topic does not mean that the topic was not discussed during consultation. For example, we have material from 10 of the 12 qualitative sites indicating that reflective practice was part of their consultation content. Yet, we know from the outcome section on reflective practice that the 2 sites not included in the content material saw reflective practice outcomes. Therefore, despite not having specific quotes pertaining to reflective practice as content, we see that reflective practice was a part of the consultation content.

At one home visiting program, a supervisor noted that, despite her long tenure in the field, reflective supervision was completely new to her. She credited the consultant for working on reflective supervision with her, acknowledging, "I'm getting better at it, but I wouldn't have done it without her." At another early childhood center-based program, a supervisor explained that the consultant does more than just listen, "she helps me think about [an issue] and think about myself critically." One of the consultants at an early childhood center-based program shared that "modeling how to use the reflective model with [the supervisor] was a big piece" of her focus at the program. Another consultant in an early childhood center-based program shared that she found a supervisor with whom she had been working very receptive to reflective practice and believed that the supervisor saw the benefit of stepping back and thinking through her own thoughts.

Capacity Building

Reflective practice is an essential component of building staff capacity to work more effectively with children and families. One way consultants addressed capacity building was by modeling for supervisors how to encourage critical thinking and problem solving by staff. For example, a consultant explained:

> I'll kind of scaffold the coordinator and see where she's at and then I'll model what I think she should do. [For instance], we were having supervision with a teaching team and the coordinator was telling them why this child is behaving a certain way. Then they asked me, "Well, why do you think he's doing it?" And I turned and asked the teachers, "Well, why do you think he's doing it?" Because up to that point it was just the supervisor telling them what she thought, and [I wanted to get] her to kind of pull back [from] problem solving [or] affecting the situation. And then the lead teacher presented a whole lot of thinking she had on what was going on with the child that then led to a specific action that was going to take place as a result of her assessment.

Similarly, the supervisor at that same program shared her perspective on the consultant's role in the excerpt below:

> [The consultant has] been helping me talk to teachers more, ask them questions, and supporting them in a different way, where they reflect on their teaching skills to support the children. . . . [For example], so a child demonstrates this kind of behavior. "Why does he demonstrate it? When does it occur? . . . If you know now that that triggers it, what is it that you can strategize to do with him?" And it's just those little things to make them look like critical thinkers.

Consultants at home visiting and early childhood center-based programs talked about their efforts to build the confidence of staff, another critical aspect of building their capacity to work more effectively with children and families. One consultant at a home visiting program said that part of her focus with a staff member was to hold the space for her to talk about how she felt when issues arose that were new to the home visitor and help her "work through her own developmental process as a worker." Likewise, a consultant at an early childhood center-based program said that she worked with a new teacher to build "her confidence and help her recognize that she has the capability to come up with answers on her own."

Some consultants reported that, at times, staff looked to them for solutions instead of as someone to help them reflect and discover their own solutions. An early childhood center-based consultant shared a story of a teacher who had concerns about a few students and asked the consultant for resources and strategies. The consultant said that they were able to schedule a classroom observation and a conversation, but the teacher "had to get used to me sort of more asking questions than just coming right out and telling her, 'Here's resources. Here's what you do.'" She noted that many teachers were used to receiving behavioral advice and that the capacity-building nature of the Illinois Model was very different.

Self-regulation

Self-regulation—being able to monitor, understand, and regulate one's emotions or reactions—is an important component of reflective practice. Mindful self-regulation, as it is called in the FAN approach (Gilkerson et al., 2012), is the process of attending to and managing thoughts and emotions during in an interaction. Mindful self-regulation can help providers maintain or regain a calm and engaged presence. Staff reported that the mental health consultant helped them recognize the value of self-regulation and develop strategies to put it into practice. For example, a supervisor at an early childhood center-based program stated, "I use [the consultant's] techniques. . . . Learning to just take a step back, look at the situation, and don't get worked up in the moment. Think clearly about the situation before you try to act on it." Another supervisor shared similar thoughts: "[The consultant] just helps me be calm and reflective and just able to step away and then address it in a way that it's not knee-jerk or aggressive."

Processing Work with a Reflective Lens

Part of the focus of consultation with the Illinois Model was to help the staff process struggles and challenges in their work. For example, consultants worked with staff to process their feelings after a challenging situation such as making a call to DCFS about potential child abuse or neglect cases. A home visitor recalled that after she made a hotline call, her program's consultant checked in with her. The home visitor said she knew the call had to be made but also knew she was taking actions that could remove a child from their home. She appreciated that

the consultant not only expressed concern about how the family was doing but also about how the staff member felt. The consultant, she said, spent "more time making me think about myself and how I feel about something." Similarly, a consultant at an early childhood center-based program said that they helped the teachers process hotline calls by supporting their judgment.

At multiple early childhood programs, consultants addressed staff members' issues with their own anxiety in relation to their work with children and families. At one school, several staff members disclosed anxiety issues, and the consultant spoke with them about how they manage their anxiety, how it plays out in the classroom, and how they might assist children who may be feeling anxious. At another program, the consultant felt that a particular staff member was unaware of how her own issues manifested in the classroom. The consultant explained,

> [The teacher] really doesn't know how to treat children in a trauma informed way. . . . She does not realize the impact that her own behaviors and the tone of her voice and her own anxiety have on her classroom. And that's something that we've really been focused on in consulting.

At a home visiting program, the consultant helped staff process chaotic visits and how to gauge success with families. That consultant explained that an important part of consultation providing a space for home visitors to share what the visit was like, what it felt like, and consider how the home visitor could support the family. At the time of that interview, the home visitors were not familiar with the FAN, but the consultant said that she introduced the principles and strategies of mindful self-regulation to help the home visitors remain grounded during chaotic visits. That same consultant also made a point of speaking with the home visitors about understanding the families' situation.

> And then [the home visitor] was able to have some more understanding of like, she's [mom] trying to work, she's got three kids, she's on her own, like how exhausting to think about having to do all of that and maybe we can come from a place first of empathy. "Let's just hear this story for a bit and be there and try to understand what their life's like." And be reflective. . . because she's able to sit back and try to really think about [the situation] rather than concluding, "Oh, this looks terrible; we're probably going to have to hotline this too."

Working with Children and Families: Early Childhood Center-based Programs

> I think it's sort of helping teachers to think about what does it mean to support these kids no matter what they come in with. —*Consultant*

This consultant's quote captures one purpose of mental health consultation: to help teachers understand a child's family context and support their work with the whole child. In this section,

we discuss several subthemes related to consultants' work with the early childhood center-based programs: 1) considering and understanding child and family issues; 2) communication with parents or guardians; 3) classroom observations and feedback; and 4) strategies for working with children.

Understanding the Context for Children's Behaviors

At several of the early childhood center-based programs, consultants worked to help staff shift and broaden their thinking about a child who may present with challenging behaviors to consider all of the influences on a child's behavior. For example:

> One pattern that I noticed was oftentimes when there would be kid that had really serious behavioral problems or developmental delays, we would start talking about their family. And I would find out that they were in foster care or a parent had left them or abandoned them. And that was kind of a pattern of something that I found myself talking about with the teachers to help them recognize how the family structure and the dysfunction within the family impact the child's behaviors and then how to respond. —*Consultant*

Consultants also provided guidance when children were experiencing loss, for instance, the death of a parent or grandparent, and other traumas. In one case, a consultant was aware of a family in the community who was experiencing domestic violence. The consultant connected the family with the program, helped to get the child enrolled and spent, in her words, "A lot of time meeting with the preschool teachers regarding this kid and this family."

Consultants also addressed a range of developmental issues, including toilet training, night terrors, special needs, emotional development, and stages of readiness to be in the classroom. For example, one consultant found that some teachers were challenged by children with high needs who they thought should be in a self-contained special education classroom. The consultant processed this issue with the teachers, validating their feelings and discussing the pros and cons of different settings for the children, including the views of the administration. "I know from the administrators' perspective, having kids move is not always the answer," a consultant explained. "That's going to be a real challenge [for teachers] being able to see that."

In their interviews, the staff and consultants did not discuss ways in which consultation addressed racial and cultural awareness and sensitivity, and how such awareness is a critical part of understanding the child's context in relation to the program environment. However, at one program, the supervisor reported that she and the consultant discussed the supervisor's concerns that the program served a culturally diverse population but had a white, monolingual teaching staff. The supervisor explained,

> I talked to [the consultant] about my feeling that a lot of the teachers don't really understand the environments that these children are coming from. And are really having a hard time meeting families where they are. . . . I feel like the teachers have a lot of bias about families, about the different cultures of the students and things like that. —*Supervisor*

Communication with Parents/Guardians

Consultants addressed communication with parents and guardians at many of the programs. Because the Illinois Model focuses on building the capacity of staff, it encourages consultants to support staff in communicating with parents rather than to meet with parents without staff present. In some instances, nonetheless, a director or supervisor asked a consultant to meet with parents directly, and the consultant complied. In other instances, consultants participated in meetings with staff and parents. When staff were preparing to meet with parents, consultants also discussed ideas for what and how to communicate with parents, especially with difficult issues like domestic violence. At one program, there were concerns that a child might be struggling with anxiety and sensory issues. According to a supervisor, the consultant and the teacher talked about the concerns and then the consultant "helped that teacher meet with the parent of the child to just be that backup person."

A number of supervisors, in fact, frequently reported that a frequent activity of consultants was helping teachers strategize to communicate with parents. For example, one supervisor said:

> [The consultant's] thing was for the teachers and the parent and all of us to be on the same page. So, if your redirection [of a child's behavior] at home is this and our redirection at school is this, let's try to find a common ground where the rules are going to be the same. So she [consultant] had meetings with teachers, admin staff, and the parent so we can all be on the same page. So, [the consultant] was, in my eyes, opening up the relationship where it was more of the parent and the teacher talking more with each other for the [benefit of the] child. —*Supervisor*

At another program, a teacher was not aware of what a child was experiencing outside of the classroom. Through consultation the teacher became motivated to "ask the dad how the kid's day started," which helped to bridge communication between school and home and add insight into the child's set of circumstances.

Understanding Classroom Contexts

As noted in the consultant logs, early childhood center-based consultants spent much of their time observing classrooms. At the start of implementation, the consultant observed classrooms to become familiar with a program and its staff. Later, a specific child's behavior challenge

frequently precipitated observations and follow-up discussions to help teachers develop strategies to respond effectively to the challenge. One consultant referred to the process as "coaching in the moment." Another consultant reported raising issues noticed during observations that might be able to be addressed by changes in the teachers' approaches. For example:

> If I'm in a classroom and the teachers are wanting the 2-year-olds to stay at tables and just sit for a bulk of time, or they're trying to transition to our gym or our outside area where kids are eager to be active and teachers are saying, "Well, you got to hold still and show me you're ready," those kinds of things, then I will work on bringing those up in ways that we can talk about it.
> —*Consultant*

In another instance, a director shared that an experienced teacher who rarely sought help did so after a new child shouted at her, tried to hit her, and tried to run out of the building. After observing, the consultant helped the teacher "create some different avenues for [the child] to release those emotions [and] talk about the emotions." Another director shared that after observing a child with academic delays and behavioral issues, the consultant was able to help the staff "look at things from a different perspective." Such examples were prevalent in the interviews.

In an extreme example, a director reported that teachers were concerned about a child and requested that the consultant conduct an observation. As a result of the consultant's observations and feedback, they were able to meet with the parent and learned that there was an incident for which the child had been in therapy but had stopped. They were then able to work with the family. The director noted, "I think that was one of the highlights that I felt like I don't know if we would have been able to do it without [MHC's] observation, you know.

Strategies for Working with Children

The interviews were full of examples of particular strategies developed during consultation. A recurring theme across programs was one of consultants encouraging teachers to "pay attention, monitor, [and] document." By noticing when and why behaviors occur—for example, if they occurred more often during transitions or if a child seemed to get overwhelmed with change—they could use consultation to wonder about and arrive at ways to help the children. One consultant provided teachers with a chart to help them track students' behaviors and then reflect on what was happening throughout the day. At one program, the consultant recommended using weighted stuffed animals or weighted blankets to help children with sensory issues feel more secure. Some consultants suggested giving children more choices so that they have more control over their situations—for example, by asking a child, "Do you want to clean up right now or do you want to wait?" Or perhaps more one-on-one time was

necessary with a particular student. At one program, the consultant suggested allowing a child to have a special classroom job that was separate from the classroom jobs regularly divided among the students. By having a special job, that particular child could feel more special while contributing to the classroom environment.

Another common topic was ways to help children communicate their feelings. At one program, the consultant suggested teachers use a communication book. At another program, teachers employed emotion cards to help a child express his feelings. One teacher said:

> [Consultant] brought in some really good techniques to address his emotions, to help him calm down first, and have him express his feelings. And if he couldn't express his feelings, we could facilitate for him by showing emotions on cards. We would show him and say, "Is this how you feel? Do you feel frustrated? Do you feel sad? Do you feel upset? Do you feel sick?" Then the child would tell us, "I just feel frustrated." Then we'd say, "Why do you feel frustrated? What makes you feel this way?" He would tell us because this friend is not sharing or because this friend just took his toy. Based on that we were able to help him learn how to express his emotions. —*Teacher*

Working with Children and Families: Home Visiting Programs

In their interviews, the narratives of home visiting staff and consultants largely focused on how consultation supported home visitors' work with families. This broad category includes: 1) engaging families, building rapport and relationships; 2) identifying and understanding family issues; 3) supporting families; and 4) cultural awareness and sensitivity.

Engaging Families, Building Rapport and Relationships

Engaging families and setting home visit expectations are the first steps to building rapport and developing relationships with families receiving home visiting services. Thus, it is not surprising that consultation at home visiting programs often focused on these topics. One consultant reported discussing strategies to encourage parent/child interactions when a parent views the home visit as an opportunity to attend to something else while the home visitor is with the child. The consultant said that she and the staff worked on "how to invite parents in, or how to be clear about what the home visit looks like, or to partner with them around having those conversations." Consultants also helped staff develop strategies for having conversations with parents about difficult issues, such as a child's behaviors or a parent's mental health issues, while maintaining their relationships. Some consultants also provided presentations on these topics for all staff at team meetings.

Identifying and Understanding Family Concerns

Consultants at all of the home visiting programs worked with staff on identifying child and family issues. At one program, the consultant showed home visitors how to build family trees to illustrate familial dynamics and see how the relationships between parties contributed to the family. Consultants worked to help the staff to consider additional factors that could impact a family's functioning or circumstances. In the words of one home visitor, "[The consultant] does a really nice job of reminding us, like, let's think how this parent feels and why they may be resistant to this or acting in this way."

At another program, a consultant was concerned about the documented prevalence of substance misuse in the community. She addressed the topic with the staff to help ensure that they would be prepared to consider how substance misuse might affect the families on their caseloads. Likewise, this consultant raised the possibility of a child having fetal alcohol syndrome and discussed its potential effects on the family as well as how the home visitor could work with the family. Consultants also addressed issues such as parents' mental health, domestic violence, and immigration status.

Identifying and understanding family issues is only part of the picture. Home visiting exists to support families in achieving healthy outcomes for their children. To that end, consultants at each of the home visiting programs concentrated some of their work with home visitors on ways to support families with their concerns. For example, according to one consultant, "We had a lot of conversations about supporting families with children with special needs or families with challenging issues such as immigration. And even just things like goal planning with families and how to provide that support."

Interviews with home visitors and consultants provided numerous examples of ways consultation helped home visitors reflect on and strategize how to support families through specific circumstances. These cases included families in which a parent had untreated mental health concerns, families who did not fully comprehend a child's delay or developmental disability, and families who dealt with multiple crises. In addition, a consultant described how she worked with a home visitor who struggled with a family that had some environmental cleanliness issues. The consultant asked the home visitor how the parent thought about the impact of cleanliness on her children. That question opened up the conversation and helped the home visitor come up with strategies to address the issue as safety concerns related to the child's developmental stage.

At one program, some home visitors were hesitant to bring certain activities that were part of the curriculum to families because they did not see them as appropriate for the families. The consultant and staff discussed the benefits of the questionable activities, ideas for those

activities, and ways to talk with the families about the activities. Another way consultants helped was to support the home visiting staff during parent group meetings. A home visitor explained,

> She [the mental health consultant] has also come to help support us during our parent group encounters. . . .For example, one of the topics was guiding a child socioemotionally, and how to help shape behavior and how to encourage positive behaviors and things like that. And so she came and helped facilitate that group meeting with the parents. —*Home visitor*

Cultural Awareness and Sensitivity

The topic of cultural awareness and cultural sensitivity emerged more often in interviews with home visitors than in interviews with early childhood staff. Issues of implicit bias, cultural norms, and cultural differences were discussed as being part of consultation during interviews with almost all of the home visiting programs. Consultants across programs raised questions and wondered with home visiting staff about how issues of cultural awareness and bias could impact their work with families. For example, one consultant said she facilitated conversations about cultural sensitivity and implicit bias at team meetings by asking questions about how they engage with families, for instance, "What do we think about when meeting new families, how to enter respectively. . . and checking out biases and our stance with that?" Likewise, another consultant discussed similar questions with staff, for example, "What are our own cultural, implicit biases and beliefs? And when they don't match the families that we work with, where does that leave us in terms of supporting the family?"

A specific example offered by one consultant involved a family who ascribed to cultural norms in which women did not typically attend social outings on their own. This clashed with the goal of having the mother attend the program's parent groups and other socialization events. The consultant and staff talked about how they could approach the issue in "a way that will make her [client] feel comfortable as a wife. . . and what the conversations looks like [with the father]." They focused the conversations about the groups on the benefits for child development and parent–child relationships, with meeting other people as just a byproduct of being in the group. Framed this way, with the focus on learning more about their child, the family became more open to the idea of the mother and child attending the groups.

Another consultant described a case in which a home visitor was concerned about a child's behaviors and how the family was disciplining the child. The consultant and home visitor strategized and decided to ask the parents open-ended questions about how they were disciplined as children in their country of origin. They said this approach could help in

> opening the door to ask questions that maybe they hadn't considered to be able to take back to their work and learn more about the family and the culture.

> . . versus just saying, "Well, they're not doing this and this is why it's happening in their house." [Instead asking,] "Well, what's different about their culture or what they do?" —*Consultant*

Work Relationships

Staff and supervisors in all of the intervention programs—both early childhood centers and home visiting programs—addressed the broad category of "work relationships" during their mental health consultation. The work relationships theme included team building, staff relationships, and staff roles.

Team Building

A common area of focus during consultation was team dynamics. Several consultants conducted team-building trainings. Consultants were called upon to discuss aspects such as cohesiveness, boundaries, and group work habits—issues and relationships that impact the program as a whole. Indeed, a consultant at a center-based program, after a summer break, felt that her major focus over the next 6 months of the program year would be around "team dynamics and building this team." Likewise, a home visitor shared that the consultant with whom her program worked did "a great job at trying to get us together as a group, realize that we do have different feelings, and that my experience is different than yours." At an early childhood center-based program, a consultant facilitated a team-building meeting at which staff discussed "barriers versus bridges." Reflecting on the session, a supervisor noted that it had the positive effects of making the teachers feel more heard and strengthening relationships across the team.

Another aspect of team building facilitated by consultants at a few programs was choosing topics to cover at team meetings. At one program, the supervisors and directors used some of their consultation time to brainstorm issues that would be helpful to cover at the team meetings. For example, if staff seemed frustrated with certain aspects of their work, supervisors could bring their concerns to the team meeting with the consultant helping to process the issue and get feedback.

Reflecting on the environment in which they work is also part of team building. At one of the programs located in a community that staff described as experiencing a lot of violence, the consultant spoke with the staff about the levels of community violence and how it affects their work. As the consultant explained:

> Violence is part of their community reality: It's [raising] some awareness. In the last two, three months, we have had drive-by shootings and the workers are out there, maybe down the block or close by. And they seem to be unaffected by those things. And I just pointed out to them, saying, "Hey, I just can't help but

notice that you talk about it like it's part of the regular day," and they said, "Well, sometimes, it is part of the regular day." —*Consultant*

Staff Relationships

In addition to working with the programs on bolstering team dynamics and team relationships, consultants also worked on issues and relationships among individual staff members. Consultants worked with supervisors, home visitors, and teachers to process and discuss their relationships, often during team meetings. As a consultant at an early childhood center-based program noted, "It's interesting, they're not so challenged by their families. They're more so challenged by the colleagues and the dynamics with their staff and with their boss, so that's where the work has been centered." Another consultant echoed this, saying, "a big part of what they have to keep processing is how they're getting along with this employee or that employee." For example, at a home visiting program a consultant worked with a supervisor about her feelings supervising a particularly challenging staff member and thinking about strategies that could help her manage their relationship and this person's work.

Consultants also focused on helping teaching teams navigate their relationships with their co-teachers. For example, a consultant in an early childhood classroom helped a teacher process her anger towards her co-teacher and maintain professional standards and boundaries. The consultant explained that the teacher was caught between articulating how she felt about the position the co-teacher had put her in (the position put a lot of stress on her), while also trying not to be hurtful with her words. In another early childhood program, conflicts arose on a teaching team because of an age and experience difference between team members. One of the teachers expressed her frustrations about the power differential in the relationship and feeling undervalued. In turn, the consultant worked with each of these teachers to help the more experienced teacher hear her colleague's perspectives and allow more freedom in the classroom.

This example raises another important aspect of work relationships that consultants addressed at some programs: how to consider the perspectives of fellow staff members. As one supervisor noted: "I think I struggle more sometimes with my frustrations with my teachers and [the consultant] helps me to ground me and think about their perspective." Another supervisor said that her consultant "talks me off the ledge and gives me ideas or suggestions or will explain to me why somebody might feel that way or be that way.

Roles

Some programs received support from the consultant regarding clarity about roles and responsibilities in the classroom or home visiting program. Understandably, changing roles within a program impacts team cohesiveness and work relationships. Consultants addressed the changes in dynamics that resulted when staff changed positions within an organization; for

example, they talked with a person about how she saw herself in the new role and how to build her capacity to take on her new duties, as well as her thoughts about how she could support the person in her prior position. In some instances, a consultant had to discern if particular challenges between staff and administrators were due to personality issues or a lack of clarity about roles or program structure. Boundary issues around roles also surfaced during consultation. At one program a consultant discussed with a director how a support staff member's "weak professional boundaries" had the potential to cause her to overstep her involvement with families and cause confusion and challenges.

Additional Content Themes

In addition to the major themes addressed above, consultation also included: 1) connections to resources, 2) training, 3) grant compliance and monitoring, and 4) personal issues. These topics were not as prevalent among the programs as those discussed above, reflecting the flexibility of the Illinois Model to address the needs of individual programs.

Connections to Resources

Several programs referenced using consultation to discuss resources for the children and families with whom they work—both resources for basic material needs as well as mental health resources. This included helping to identify resources for families, as well as processing the fact that resources may not be readily available and that sometimes families opt not to use the available resources. Home visitors at a few programs shared that they met with their consultants about families who lacked resources, the impact that has on the family, and their work with the family. For example, one home visitor said:

> Our job is to not fix the family's problem. Our job is to give the family resources. So, she'll [consultant] help provide, if it's something we are struggling with getting a resource for, she'll help come up with ideas for resources. I think that's probably the biggest thing, dealing with those challenging behaviors. Just knowing that we don't have to fix them, but we can support them. —*Home visitor*

A home visitor at another program shared similar thoughts: "And so we have a lot of conversations about resources available in the area that can be utilized that can support the family. . . . That's another big thing we talk about quite a bit." Early childhood center-based programs also referenced consultation as an avenue for discussing resources. A teacher at one program said that the consultant introduced them to a local program that works with children exposed to violence and was glad to learn of that resource, which could potentially benefit many students. She noted, "[Consultant] was able to kind of make those connections for us as a resource that can help us improve."

Training

Consultants facilitated periodic trainings as part of their work for this initiative. However, these were not described at length in the interviews. Training topics referenced during interviews include impact of trauma; having difficult conversations with parents; brain building; maternal mental health, postpartum depression, and reflective supervision; and sensory processing. Consultants also addressed self-care strategies with staff.

Grant Compliance/Monitoring

Issues related to grant compliance and monitoring as part of consultation content were raised during interviews with a few of the early childhood programs. At one program, the consultant noted that teachers were coming to her stressed about various accreditations and monitoring issues. She recalled a teacher who was worried about how to navigate what felt right to her with regard to her interactions with her students with what she was supposed to do to follow rules and regulations in advance of a monitoring visit. The teacher felt that she could not "really be available" to the children because she would be so focused on what someone who came to observe her room would see.

At another program, a new administrator realized the program was out of compliance in some areas of its curriculum model. The consultant worked with the supervisor to message the necessary changes for the staff to help bridge the divide between what they had been doing and how their practice had to shift. Strategies included helping to explain why the changes were necessary to comply with grant, acknowledging that change is not easy, and that implementing the changes would benefit the children with whom they work. In conjunction, the consultant worked with the supervisor and teachers to process the changes and move towards compliance.

Personal Issues

Sometimes consultation focused on the staff's personal issues. For example, a consultant said that she talked with staff about their family needs, like the sickness or needs of their own children, which could affect their work. She found such conversations to be "a nice little opportunity to address the likelihood that a tired mom is a tired teacher. [And it was] nice to talk about the strength building in both roles, 'cause if you're not good at the home one, the one at school is going to suffer and vice versa." At another program, the consultant noted that some of the staff "enjoyed talking about things that were stressful for them in their personal environment." Some staff spoke with their consultant about professional development and processing options in their careers.

The challenge was to bring the topic back to how personal issues can affect the work of the staff. Some consultants were better at bringing the discussion back to this than others. At a few programs, consultants struggled to keep personal issues from dominating a consultation

session. As one consultant explained, especially at the beginning of her relationship with a program, she let the staff "decide what they need to talk about" and "sometimes, when they're dysregulated, we end up talking about their own personal struggles a little bit more." She acknowledged that the discussion of personal issues was "a little bit more than I'm actually comfortable with." She worried that her consultation was not focused on what it should be, although she said she tried to refocus the group too. The supervisor and director at this program expressed similar concerns, suggesting that the group sessions needed to have more structure and reflection and less counseling for personal issues. According to another staff member, staff talked about their own issues "because they thought that's what it [consultation] was."

Diversity, Equity, and Inclusion Issues

One of the core competencies for mental health consultants is the ability to work effectively among diverse cultures and communities. As mentioned in the previous section, the issue of cultural sensitivity was raised by a few consultants and supervisors during their interviews, but, overall, it did not appear that discussion of this topic occurred frequently during consultation with program staff and administrators.

This competency has assumed heightened importance in conversations about suspensions and expulsions in early childhood programs, which have historically disadvantaged children of color. As part of their training and ongoing supervision, consultants engaged in numerous discussions about diversity, equity, and inclusion (DEI) issues with other consultants and experts in the field. One of the Illinois annual consultant retreats was devoted to this topic. Consultants also participated in an all-day professional development workshop on "Diversity-Informed Infant Mental Health Tenets" (Tenets Initiative, 2018) developed by the Irving Harris Foundation Professional Development Network in 2012 and updated in 2018. Given the emphasis in the pilot on DEI, we were interested in the extent to which DEI issues were encountered during the implementation period and which tenets were incorporated into the consultation process.

Consultant Incorporation of DEI Concepts in Work with Staff

Consultants all shared generally positive feedback about the Tenets trainings and expressed the view that it is important to "do DEI work" both personally and with program staff. However, most also explained that they had not fully internalized the Tenets or actively applied them to their work with program staff. For example, a consultant explained, "So I don't remember a ton of the training. I mean I know it's always helpful. I'm not probably doing a really wonderful job of like taking it to the next level." Similarly, other consultants expressed, "I guess I haven't felt like I have really used it," and "I could do a better job." Another consultant explained how the training had been helpful for beginning to notice racial disparities at the early childhood program, but that there was a need for ongoing work:

> I do enjoy that training. But I need tons more of it. Tons more. But it has helped me, obviously, because I'm seeing more, you know, the leadership's all white. These people are all black. And this doesn't make sense. I'm kind of just seeing things a little bit differently. I mean, it has been helpful. But that's going to be an ongoing thing. I mean, one training is not going to do it at all. —*Consultant*

Consultants also noted that it was challenging to find opportunities to have these sensitive and uncomfortable conversations with program staff. For example, another consultant explained:

> I actually have gone to about three of [the Tenets trainings] I think it is great, a great discussion. I think it's great to have those written in front of people. I just don't feel like it is—I feel like it's a discussion that in the consultation world, we're having a lot, but it's not happening as often or even as efficient maybe in a lot of programs. And so it's just the idea, I just always wonder, is it—like I said, with my program, kind of pointing out what culture is or what race is or what ethnicity is for us and what it means to our work has to be a topic that's really intentionally presented and just sometimes because there's so many other things going on, you are not always sure, like, is this a good time to bring it up. And you want to wait for everyone to be comfortable, which may not happen. —*Consultant*

Self-Awareness

Consultants varied a great deal in their discussions of their own self-awareness and their comfort with having DEI-related conversations with program staff. While most consultants noticed DEI issues at the programs where they were placed, most program directors and supervisors reported that they did not have these kinds of issues at their programs. This discrepancy suggests a general lack of awareness or discomfort speaking about these issues among program leadership. Some consultants highlighted the importance of raising self-awareness of implicit biases in order to improve relationships among program staff. One consultant explained that part of the consultant role was to help raise this awareness. "So, how do I create the awareness that these are factors that we deal with because of our work?" However, complex relationships and power dynamics can make it very challenging to have such sensitive conversations, especially with program directors and supervisors.

Advocating for Children and Families

In their discussion of suspension and expulsion policies, supervisors explained that their programs no longer permit suspending or expelling pre-K students. While some respondents noted this was due to changes in the law or Head Start policy, others indicated it was due to program values. Some directors and supervisors noted that there were teachers who were frustrated by their inability to send a child home or to a different classroom. Additionally, some

programs explained that they may ask for help from an administrator, other staff, or even a parent to assist the child within the classroom or take the child out of the classroom for a break. For example, one teacher explained that due to the lack of space and support to calm a child at her school, she has sent children to the office against school policy:

> We are told we're not to send anyone to the office for any reason or anything, but I did. There were times where I'm like, "I need a break." I physically am exhausted, I'm mentally exhausted. You're trying to stay calm, you're trying to do the strategies and everything. One of the strategies that we should be able to do is take them in the hallway—you know, for these particular students they needed a physical outlet, whether it's just throw a ball on a wall or take a walk or do anything. Because of where we are located, we are not allowed to do that. There should not be action or movement in the hallways. . . . And we do not have a place for these students. A lot of other schools have rooms where it's got like the bouncy balls and a padded room where they can release some things. We do not have anything like that. And because it is just me I can't leave the classroom to take them to the playground or outside of the school. You think about all the different strategies, like well how do I do this, how do I work it? I know this is what this child needs but how do I work that within what's available? —*Teacher*

Staff from two programs also said that if behavior issues did not de-escalate, they would occasionally send a child home early. Directors and supervisors explained that if a child was acting aggressively towards others, they would try to work with the family to get to the root of the issue in order to keep the child enrolled. However, there were instances in which this led to looking for a more appropriate placement or the family's removal of their child from the program. A supervisor also noted that in one instance when a child had been expelled in a previous year, a language barrier with the family made it challenging to discuss and address the child's behavior issues.

Only one consultant brought up the issue of racial bias influencing teachers' mandatory reporting of child abuse and neglect; however, this is a complex issue worthy of further attention. In this instance, the consultant was concerned that a teacher's fear of retaliation from black families prevented her from making mandatory reports to Child Protective Services (CPS). Importantly, this same bias could also lead a teacher to make a CPS report when one is not actually warranted. In both instances, the consultant's advice could help mitigate the risk of allowing implicit bias to influence a teacher's decision to report child abuse or neglect. As the consultant noted,

> The primary focus was on reminding her of the law regarding mandated reporting, and what we are and are not required to do. How that works. Who

she should talk to when she thinks she needs to report something. How to seek assistance and like emotional support when she does that, if she's had concerns. —*Consultant*

Advocating for children includes ensuring children with special needs are appropriately supported and included in classroom activities. Consultants explained that some teachers found it challenging to integrate children with special needs into their classrooms. One consultant also discussed how a program's bureaucratic transportation policies created a barrier to attendance for one student who was severely physically disabled.

Increasing Awareness of Inequities in Systems

Many of the program leaders and staff, and even a few consultants, seemed to have narrow concepts of diversity, seeing it as primarily referring to race. For this reason, many expressed the view that because of the racial homogeneity or racial diversity within their program, there were no DEI issues to address. However, this perspective fails to acknowledge the ways privilege and discrimination can operate within both homogenous and diverse populations.

Additionally, one consultant noted how she believed that teachers' implicit biases led them to perceive more challenges with the black boys in their classrooms:

> It would be something that would be good to have more open dialogue and conversation about. I do notice that the kids that are most brought to my attention tend to be the older African American boys in classrooms and it certainly fits with much of that implicit bias work. —*Consultant*

Respecting Nondominant Bodies of Knowledge

As mentioned above, issues of implicit bias, cultural norms, and cultural differences were more often discussed during consultation with staff and supervisors at home visiting programs than in the early childhood center-based programs. Consultants discussed the importance of approaching families in a culturally sensitive and respectful way. For example, a consultant explained how asking reflective questions was a helpful approach for adjusting home visitors' perspectives about serving families from different cultures:

> I think it's that idea of asking the open-ended and reflective questions. One of the things we talked about with this worker and that family was about, she was concerned about the daughter's behaviors and about how the family was addressing discipline issues. So, we had a conversation of opening it up. Like "Have you ever talked to them about what discipline was like for them in [their home country]?" She was like, "Oh, I haven't even thought of that." So, it's just opening the door to ask questions that maybe they hadn't considered to be

able to take back to their work and learn more about the family and the culture. —*Consultant*

Support Families in Their Preferred Language

Several interview respondents noted that their programs offered bilingual education and supports, ranging from dual language programs to ESL programs. However, some of these programs also reported a shortage of Spanish-speaking staff that prevented them from fully implementing bilingual programming. Spanish was typically the second language, and programs noted language barriers for communicating with or providing referrals for supportive services to families who spoke other languages.

Allocate Resources to Systems Change

Consultants reported that it was challenging to find the appropriate time and space for sensitive and uncomfortable DEI conversations. This can be especially challenging for programs with leadership that does not recognize the relevance and importance of these issues. For example, one consultant explained, "But there honestly is not a lot of opportunity to talk about diversity." Another consultant described her efforts to advocate for the program to bring someone in to lead conversations about diversity, equity, and inclusion.

Make Space and Open Pathways for Diverse Professionals

Interview respondents discussed how, in some cases, consultants, program administrators, and teachers were not representative of the families they served. Consultants discussed ways that they advocated for the inclusion of more diverse teams, but also noted the limitations of their role in pushing for this goal. One consultant said, "So I'm trying to give hints and just, without coming out directly and say, 'You know, be nice if we had some diversity in the leadership team. I think that's what we're missing for this communication. How can we do that?' 'Cause I don't have the relationship with them."

A program director acknowledged that they would like to have more diversity within their leadership team but have not had a diverse applicant pool that meets the qualifications they are looking for. At the same time, the director's remarks below indicate a willingness to think differently about recruitment and hiring for leadership positions:

> I think as an agency, we are always looking to diversify our management team. I think that's where some diversity is needed, but we also recognize that we hire the most qualified candidates. While we're aware that maybe our management team doesn't necessarily look like the families we serve, we have to hire the best qualified person no matter what they look like. That has been our commitment and that will continue to be our commitment, but we're always aware of things that maybe we can do differently. —*Program director*

Additionally, a supervisor noted the importance of not only making space and opening pathways for diverse professionals, but also ensuring fair pay. The supervisor said, "I think that one issue that we had with the equity, making sure that everybody is, for one, paid the same for the work that they are doing, and those type of things. So, that's something that the agency looks at, as a whole."

Views of Mental Health Consultation

The qualitative interview data provide numerous examples of the issues that were discussed during consultation and the many ways that consultation benefited staff and supervisors. At baseline and at each subsequent data collection point, staff and supervisors were asked in their online surveys to rate their experiences with and views of mental health consultation. Table 5 shows ratings of consultation by staff and supervisors in the intervention group at baseline for the analytic sample. It should be noted that only those who reported having access to a consultant during the previous 6 months responded to the questions about the quality and value of consultation, so sample sizes are rather small, and differences should be interpreted cautiously.[23] Overall, responses were similar between the staff and supervisors except for their ratings of the value of consultation; supervisors tended to rate consultation as "very valuable" whereas staff rated it as "moderately valuable" at Time 4.

Staff in the intervention group who responded to the survey tended to rate the value of the consultation at Time 2 and Time 3 somewhat lower than at baseline and Time 4. This might reflect the adjustment staff were making to the new approach to IECMHC as well as the time it took to build staff reflective capacity and understanding of the Illinois Model. In comparison, supervisors' ratings were more even across the post-implementation periods. In other ratings, the staff and supervisor responses were similar. Both staff and supervisors agreed that the amount of consultation they received was adequate to meet their needs, but that scheduling consultation was not always easy. In terms of overall quality, most of the ratings fell between "good" and "very good." There was a small decline in supervisors' ratings, which might reflect their perspectives during the intermittent support period or transition to that period.

[23] We do not know why some staff in the intervention group did not respond "yes" to having access to consultation. Perhaps some staff were not aware of the consultant or did not know the role of the consultant (possibly referring to them as a different title). In addition, this measure was administered to staff in the comparison group, but we do not include their responses here because we are following an intent-to-treat approach in the analysis. We also have limited data on the nature of the consultation received by the comparison group, and it would be difficult to interpret any differences or lack of differences we obtained without more information about the comparison group.

Table 5. Views of Consultation over Time by Staff and Supervisors in the Intervention Group

Indicator	Staff (*N* = 72)				Supervisors (*N* = 14)			
	Baseline *n* = 39[a]	Time 2 *n* = 50[a]	Time 3 *n* = 51[a]	Time 4 *n* = 36[a]	Baseline *n* = 10[a]	Time 2 *n* = 12[a]	Time 3 *n* = 11[a]	Time 4 *n* = 7[a]
Value of consultation[b]								
Mean (*SD*)	3.2 (1.14)	2.6 (1.19)	2.7 (1.17)	3.0 (1.06)	3.5 (0.85)	3.6 (0.93)	3.4 (0.81)	3.9 (0.38)
Range	0–4	0–4	0–4	1–4	2–4	1–4	2–4	3–4
Ease of scheduling[c]								
Mean (*SD*)	2.5 (1.19)	2.3 (1.38)	2.6 (1.19)	2.3 (1.15)	2.4 (1.26)	2.7 (1.10)	2.7 (0.79)	2.4 (1.72)
Range	0–4	0–4	0–4	0–4	1–4	0–4	2–4	0–4
Quality[d]								
Mean (*SD*)	2.7 (1.24)	2.5 (1.03)	2.6 (1.00)	2.8 (1.00)	2.7 (1.16)	3.3 (0.65)	3.1 (1.04)	4.0 (0)
Range	0–4	0–4	1–4	1–4	1–4	2–4	1–4	4–4
Adequacy to meet needs[e]								
Mean (*SD*)	1.3 (0.45)	1.2 (0.39)	1.1 (0.33)	1.1 (0.32)	1.3 (0.48)	1.0 (0)	1.0 (0)	1.3 (0.49)
Range	1–2	1–2	1–2	1–2	1–2	1–1	1–1	1–2

[a] Only those who reported having access to a consultant during the previous 6 months responded to the questions about the value, quality, etc. of consultation, so we indicate the range of sample sizes for these items.
[b] Response scale: 0, "Not at all valuable"; 1, "A little"; 2, "Somewhat"; 3, "Moderately"; and 4, "Very valuable."
[c] Response scale: 0, "Not at all easy"; 1, "A little"; 2, "Somewhat"; 3, "Moderately"; and 4, "Very easy."
[d] Response scale: 1, "Yes, the amount of consultation from a mental health consultant is adequate"; 2, "No, the amount of consultation from a mental health consultant is NOT adequate."
[e] Response scale: 0, Poor"; 1, "Fair"; 2, "Good"; 3, "Very good"; and 4, "Excellent."

Implementation Factors

In their qualitative interviews, intervention program staff and consultants identified a number of factors—facilitators and barriers—that affected implementation of the Illinois Model. These factors are organized into three main themes: 1) leadership and staff buy-in; 2) understanding the Illinois Model implementation and goals; and 3) structural or administrative factors. There were also a few other factors, which typically only occurred in one or two sites. The last category is included because it provides insight into potential barriers to implementing the model.

All the intervention sites and consultants reported some implementation barriers. The extent to which those barriers remained throughout the duration of the pilot, however, varied. For many programs and consultants, initial barriers—particularly lack of leadership and staff buy-in—were less of a concern as time went on and as consultants and staff developed relationships and trust. For two programs, the initial barriers remained in place and implementation was challenging throughout the duration of the pilot.

Leadership and Staff Buy-in/Readiness for Consultation

The first two main themes—leadership and staff readiness for and understanding of MHC implementation and goals—were closely related but distinct. Most of the programs and consultants reported some challenges related to lack of either leadership or staff buy-in. At times, the lack of buy-in and lack of understanding of MHC implementation and goals were intertwined: leadership and staff did not understand the purpose of consultation and therefore did not support it as a valuable endeavor.

Leadership Buy-in

Leadership buy-in was instrumental in a successful implementation, as it informed how much of a priority the consultant's work was at the program, and leadership made decisions regarding making time for reflective supervision. At times lack of leadership buy-in was related to leadership turnover. In a few programs, program leaders enthusiastically signed on to receive consultation, but early on in the implementation process, they left or retired from their positions and were replaced by individuals who were not familiar with the IECMHC concept and/or had other priorities as they assumed their new positions. This created confusion and a diminished the importance of consultation for supervisors and staff; and made it challenging for the consultant to work effectively and for the staff to have guidance as to how to respond to consultation.

Staff Buy-in

In many—but not all—cases, lack of staff buy-in coincided with a lack of leadership buy-in. However, in some programs, staff were reluctant to involve someone new in their process despite leadership enthusiasm. For some, the perception that consultation consisted of discussing emotions and requiring vulnerability, particularly in light of their numerous demands on the job, contributed to staff's reluctance to engage. For example, an administrator recalled one of the frontline staff's initial resistance to working with the consultant: "[The staff person] was like, 'I'm not doing this. I'm not talking about my feelings. I'm not going to tell you what I think. I don't have time for this.'"

Understanding MHC Implementation and Goals

About half of the programs and consultants reported an initial lack of clarity regarding implementation and goals of the Illinois Model among staff. In many cases, they came to understand the consultant's purpose enough that it was no longer a barrier. In a few cases, however, staff never really understood the consultant's purpose and thus did not fully engage in consultation.

Just as the program staff were not always clear on the goals of the model, at times the consultant did not add sufficient clarification. For example, a supervisor commented:

> I almost wish there was more of a structure to the meeting. . . . I feel like it's turning into a chitchat session, which is good to an extent. I feel like it needs to be a little bit more focused on why we're coming together. . . . I don't think any of us really know what this should look like, these hours that we meet every week, because it's all new to us. —*Supervisor*

A consultant at another program struggled with understanding the model, and that confusion contributed to staff not really realizing they were even receiving consultation. During an interview, a teacher acknowledged, "To tell you the truth, I have heard [infant early childhood mental health consultation] from you [researchers]. I never heard of that before, and when they told us you guys were coming that last year, I felt like we didn't have enough information, so I didn't really understand what it meant, or what was it that you guys were doing."

Structural/Administrative Issues

All but one of the intervention programs reported logistical or administrative barriers. Most of these barriers involved: 1) scheduling difficulties and staff availability and 2) funding or contract issues. Funding and contract issues tended to resolve themselves without a long-term impact on implementation, while scheduling difficulties and staff availability was an ongoing issue for most sites.

Scheduling Difficulties and Staff Availability

The interviews made clear that the study programs had very hard-working staff with many obligations in addition to their work with children and families. For some programs, engaging with the consultant felt like one more thing on a long list of things to do. One staff person said:

> I've got to be real honest, when [the consultant] first got here, my very first conversation was, "This is a waste of my time. . . . " I had other things to do, time-wise. And I just saw it as one more thing that just interfered with what we needed to be [doing for children and families]. . . . I was like, "Oh, we've got to stop to talk to somebody?" —*Teacher*

This person went on to talk about how they eventually found the consultant to be very helpful, and to be someone who would listen without judgment and without giving more tasks for the staff to do. While some programs seemed to share that sentiment—after some time, they came to value the consultant and subsequently found ways to make consultation a priority— others struggled with scheduling difficulties and their impact on consultation in an ongoing way.

Some of the scheduling difficulties and lack of staff availability was due to high staff turnover. Turnover affected schedules and staff availability. As supervisors and directors worked to arrange for substitutes or fill in for missing staff, they were less available to meet with the consultants. There also were challenges scheduling time in programs associated with school

districts—whether they were early childhood center-based programs or home visiting programs—because of having to adhere to the school district's schedule. School district-associated programs were limited to a particular daily schedule (e.g., 8:30 a.m. to 3 p.m.) and that made it difficult to find time to meet with staff. In addition to having more limited hours during the school year, school-based programs often closed during the summer.

Funding or Contract Issues

Some of the programs had initial funding issues or contract complications that prevented the consultant from starting their work. These types of issues contributed to a "rocky start" for some consultants as they were not able to be present at the programs, resulting in some confusion and frustration. While these issues tended to resolve themselves, they initially served as a barrier to implementation.

Other, Less Common Barriers

Each intervention program was unique and any number of factors contributed to implementation challenges. Some of these challenges got resolved quickly while others were more lasting. Barriers that occurred at one or two programs included 1) a combination of interpersonal and structural issues and 2) consultants' lack of experience or understanding of the model.

At one program, individuals in leadership roles had interpersonal conflict, which contributed to a lack of clarity about their roles. At times, the consultant got drawn into this conflict, which interfered with implementation and dominated the consultant's time. Additionally, the lack of clarity about the leaders' roles meant that it was not always clear who the consultant's main contact or point of access to the program was.

While most implementation barriers were programmatic, a couple of the programs had inexperienced consultants or consultants who underestimated the amount of time required for implementation. Some of the confusion about the Illinois Model could be tied to the consultant's difficulties either understanding the model, maintaining boundaries, or other interpersonal issues with staff. According to a site leader:

> [The] lines were a little blurry of what [the consultant's] roles [were]. . . . The boundaries and expectations seemed a little blurred when [the consultant] was here. . . . [The consultant], at a couple of different points, gave feedback. . . [and] told a coordinator some unprofessional advice of what to do that was not really in the scope of [the] mental health consultant role. —*Program director*

Implementation Facilitators

The most important facilitator in successfully implementing consultation in the programs was leadership support and buy-in. For example, a supervisor told us that she wanted to participate in the pilot because the staff were seeing "children with challenges" that the staff found overwhelming. When the supervisors and directors were excited about the pilot, that excitement was generally picked up on by frontline staff. Further, leadership support meant that consultation was a priority; for example, schedules were adjusted to ensure meetings with the consultant occurred. In the case below, teachers turned to their supervisor when they were still building trust with the consultant. Their supervisor's support and confidence in the consultant facilitated the teachers' willingness to engage with the consultant.

> [We] had a case in the beginning. . . and it was very touchy, very delicate. So [the teachers] were like should we say it in front of [the MHC]? I'm like yeah, that's why we're here, more than anything. Cause if we offer [MHC] services to support the family. . . we need to let [MHC] know what's going on. They are like ok then, so when are we going to have supervision again? . . . So it started slowly, but then now they ask is [MHC] coming? —*Supervisor*

Additional facilitators included having regular team meetings already in place, having a prior relationship with the consultant and trust already established; and staff who were open to learning and reflecting. Other facilitators included previous experience with reflective supervision on the part of individuals and a commitment on the part of program directors and supervisors to maintain a schedule for supervision. Home visiting program staff were somewhat more likely than staff at early childhood centers to have experience with reflective supervision. However, despite this difference, both types of programs struggled to make time for regular supervision.

Chapter Summary

With few exceptions, the Illinois Model was successfully implemented in a variety of early childhood center-based and home visiting programs. We determined the success of its implementation by both structural and process indicators of fidelity. Dosage data, one of the primary structural indicators, indicated that nearly all the programs received at least 80% of their expected consultant goal hours. Consultant log data allowed us to examine adherence to the model, which is the other primary structural indicator. These data highlighted the variability in distribution of consultant activities, although all intervention programs received the expected types of consultant support.

Themes emerging from the qualitative interviews with program staff and consultants aligned with the analysis of the content of the consultant logs, which included 1) reflective practice, 2)

working with children and families, and 3) work relationships. Interview data underscored the ways in which consultants adapted their work to fit the needs of the individual programs. While consultants spoke favorably of their training in "Diversity-Informed Infant Mental Health Tenets," DEI issues were not a primary topic of discussion. DEI needs to be addressed across programs. Staff and consultant interviews also illuminated barriers and facilitators to model implementation. Barriers included: 1) lack of leadership and staff buy-in; 2) lack of understanding of the Illinois Model implementation and goals; 3) structural or administrative barriers; and 4) other, less common barriers, such as consultants' lack of experience and interpersonal conflict. The primary facilitator in successfully implementing the model was leadership and staff buy-in as reflected in providing space, modifying schedules to make time to meet with the consultant, inviting consultants to team meetings, and generally reminding staff of the availability of the consultant. Finally, staff and supervisor survey responses indicated that they found consultation to be "moderately" to "very" valuable and adequate to meet their needs. Overall, dosage, adherence, and process data indicate the successful implementation of the Illinois Model in urban and rural early childhood center-based programs and home visiting programs.

Findings: Outcomes for Providers, Children, and Families

> [The consultant] helped me realize I don't have to carry the burden, and I don't have to be a person that needs to fix everything for the families. Sometimes it's okay for me to sit in a visit and be quiet and not have all the answers. So it made me a calmer person in how I respond to certain things that come up at visits and helped me to realize ways I can reflect with the parents. Or to just let me vent [with the consultant] and then I may figure out, okay, if this happens again what could I do different or how can I minimize my reaction [next time]? So we just get to work through things a little bit more in ways that obviously I didn't or couldn't do by myself. —*Home visitor*

The Illinois Model of IECMHC was developed to improve the skills of early childhood professionals who care for and work with young children and their parents. Because program staff and supervisors received the intervention directly, the evaluation focused on understanding how to implement the model in a variety of early childhood programs. The evaluation also studied proximal outcomes—the effects of the model on staff practices and well-being. In addition, changes in staff are expected to affect children and families. Thus, the evaluation sought to understand distal outcomes—the potential effects of implementing the model on the experiences and well-being of program participants. In this chapter, we begin by presenting the results relevant to the intervention's effects on staff and then supervisors. We then describe effects on children and families.

Staff Outcomes

In this section, we report how the Illinois Model affected staff reflective capacity and well-being, staff and supervisor relationships, and staff interactions with children and families. We found significant intervention effects for staff reflective functioning, staff burnout, teachers' classroom climate, and home visitor practices. After examining descriptive statistics of our sample, we removed variables that were not statistically significant or that were highly correlated with other variables. The following variables were tested in the models predicting the staff survey outcomes:[24]

- Staff age (2 variables)
- Staff race and ethnicity (4 variables)
- Staff educational attainment (2 variables)

[24] Details on the analytic approach and the variables included in the models can be found in Appendix A.

- Program type (center or home visiting)
- Program size (small, medium, large)
- Intervention or comparison group
- Dosage of intervention
- Time

Appendix Table C-1 displays the descriptive statistics for all survey scales administered to the staff. We explain the findings for each of the following constructs below: 1) reflective capacity, 2) burnout, 3) depression, 4) self-efficacy, and 5) staff and supervisor relationships. Because of the number of different outcome measures used in the evaluation, we do not present details of every analysis but, rather, only those that resulted in findings that were either statistically significant or showing a tendency to be significant.

Staff Reflective Capacity

To understand how the intervention affected staff reflective capacity, we used two measures of reflective capacity for staff and one for supervisors. Both staff and supervisors received the Reflective Functioning Questionnaire (Fonagy et al., 2016) in each of their four surveys. Staff in the focused sample were administered an additional, narrative measure of reflective capacity in their baseline and Time 3 interviews.

Reflective capacity, or the ability to interpret one's own and others' mental states (Fonagy et al., 2016), has been found to improve the well-being and practice of early childhood providers (Cigala et al., 2019). We administered staff the Reflective Functioning Questionnaire (RFQ; Fonagy et al., 2016), a standardized measure of reflective functioning containing subscales assessing the certainty and uncertainty about the mental states of self and others. The Certainty subscale assesses genuine mentalizing, where a high score reflects a respondent's understanding of their own thoughts and feelings and those of others, while acknowledging that thoughts and feelings can be difficult to understand. A high score on the Uncertainty subscale reflects an almost complete lack of knowledge about mental states. Thus, high reflective functioning produces high Certainty and low Uncertainty scores. We found a significant two-way interaction between time and group for the RFQ Uncertainty subscales and a trend towards significance for the RFQ Certainty subscale. (See Table 6 for these results.) Internal consistency was high for the Certainty subscale at all four time points (Cronbach's alpha = .83, .84, .87, .87 at baseline, T2, T3, T4, respectively) but lower for the Uncertainty subscale (Cronbach's alpha = .72, .72, .82, .67 at baseline, T2, T3, T4, respectively).

Scores on the Certainty subscale ranged from around 2 or below 2 on the 3-point scale at baseline for both groups. However, staff in the intervention group tended to have a higher average score over time on the Certainty subscale than staff in the comparison group. This finding implies that staff who received the intervention tended to have more genuine

mentalizing and higher reflective functioning over time than staff in the comparison group. See Figure 6.

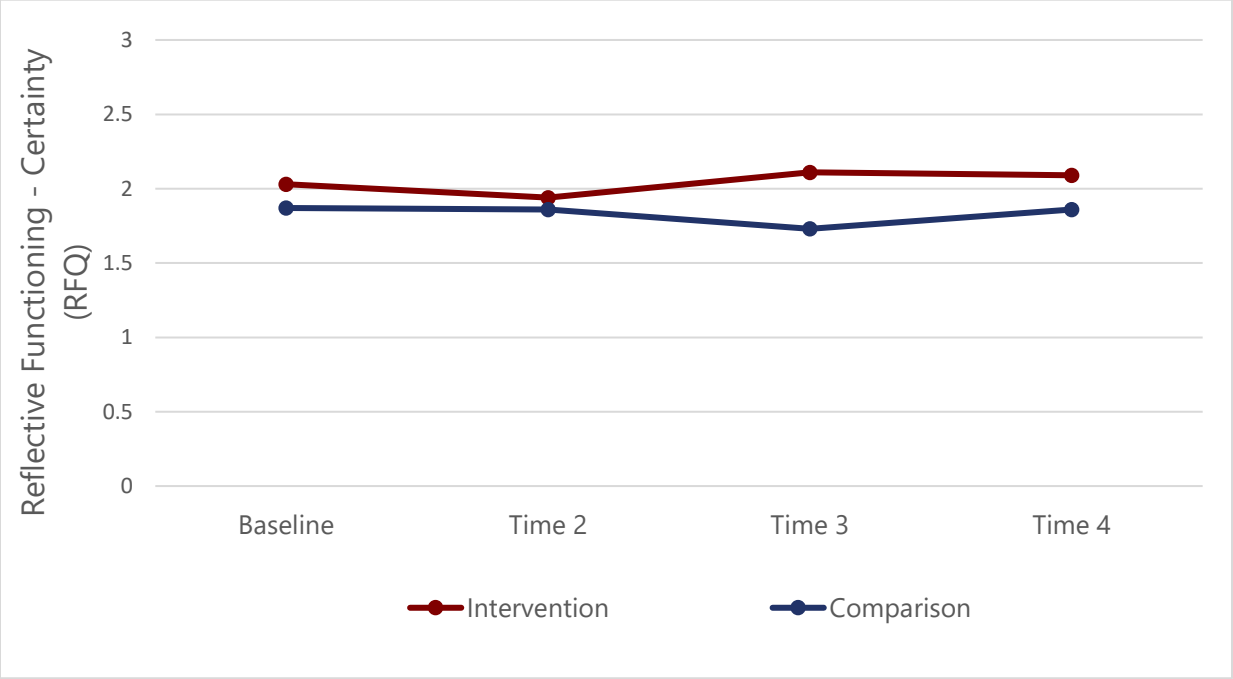
Table 6. Reflective Functioning: LMM Analysis Results (*N* = 107)

Scale	Intervention effect (Time*Group)	Effect of any other variables
Reflective Functioning Questionnaire – Certainty	β = 0.11, *p* = .068^	Staff age*Time^2, β = 0.73, *p* = .099^
Reflective Functioning Questionnaire - Uncertainty	β = -0.09, *p* = .033*	N/A

^*p* < .10, **p* < .05

Note. Our usual *p* level for statistical significance is < .05. However, we report results of *p* < .10 to note a trend towards significance in the data.

Figure 6. Reflective Functioning Certainty by Group Over Time (*N* = 108)

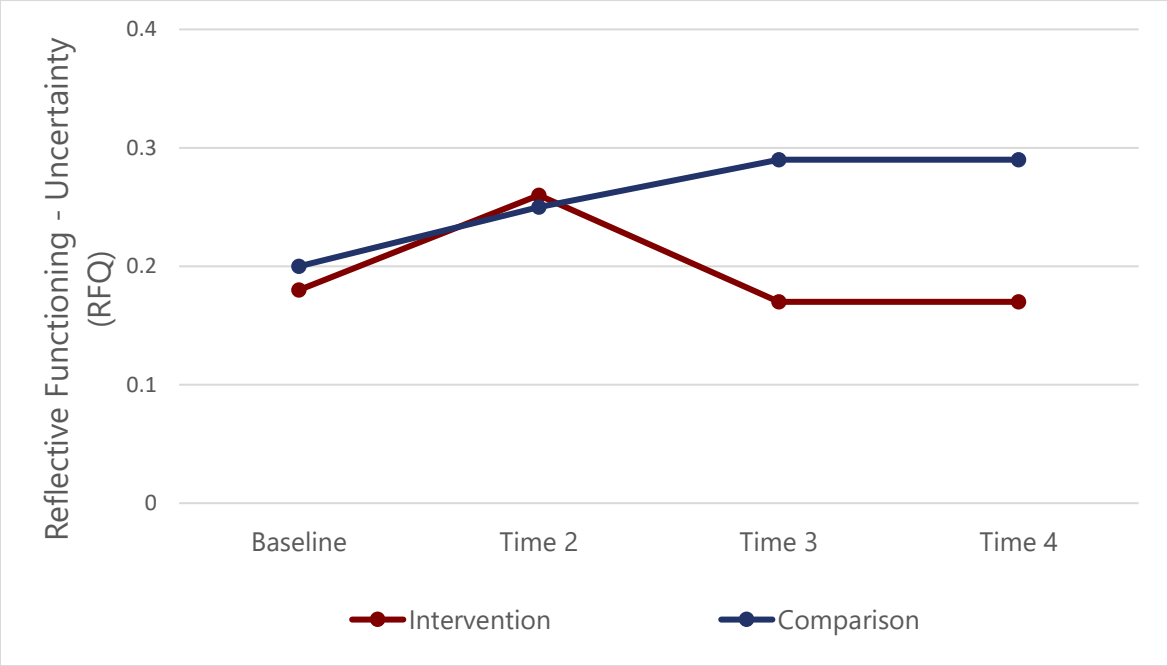


Note. Two-way interaction effect between time and group on the RFQ Certainty subscale, β = 0.11, *p* = .068. The Certainty subscale has a possible score range of 0-3. Staff scores on the Certainty subscale (*M* = 1.94) tended to resemble scores in other studies (see, for example, Fonagy et al., 2016; *M* = 1.98).

Staff age had a tendency to affect (*p* < .10) responses to the Certainty subscale over time, an interaction effect that appears to be nonlinear (see Table 6). In particular, a parabola is the best shape to describe this relationship, where Certainty increases with staff age for a period of time and then begins to decrease with staff age.

We also found a significant effect of the intervention on the Uncertainty subscale over time. Thus, staff members who received the intervention had a lower score over time on the Uncertainty subscale, demonstrating a greater understanding of knowledge about mental states, than staff in the comparison group. It also should be noted that the subscale scores can range from 0 to 2.33; thus, scores for both groups are relatively low at all time points, ranging from less than 0.2 to just under 0.3 (see Figure 7).

Figure 7. Reflective Functioning Uncertainty by Group Over Time (*N* = 107)



Note. Two-way interaction effect was significant between time and group on the RFQ Certainty subscale, β = -0.09, *p* = .033. The Uncertainty subscale has a possible score ranging from 0-2.33, with a higher score indicating higher uncertainty. Staff scores on the Uncertainty subscale (*M* = 0.23) tended to be much lower than participant scores in other studies (e.g., Fonagy et al., 2016; *M* = 1.77, yet their sample included participants from various occupations).

The other measure of staff reflective capacity was the Provider Reflective Process Assessment Scales (PRPAS; Heller, 2017). As described earlier, we administered the PRPAS with the focused sample of teachers and home visitors as part of their interviews at two time points, baseline and Time 3 (about 12 months later). Because of attrition, not all staff in the focused sample had both a baseline and Time 3 interview. Thus, our final sample for the PRPAS analysis was 13 staff in the intervention group and 13 staff in the comparison group. There were no demographic differences between these two groups and the larger sample, with one exception; the PRPAS sample had higher education levels than the analytic sample (more likely to have a bachelor's degree or higher in the subsample), chi square = 5.39, *p* = .02.

The PRPAS consists of six domains that correspond to different dimensions of reflective capacity: Self-Knowledge, Self-Regulation, Multiple Perspectives, Collaboration, Process, and Authentic

Attitude. We conducted a repeated measures ANOVA to test for differences in reflective capacity over time between the intervention and comparison groups. Results of the analyses are displayed in Table 7.

Table 7. Provider Reflective Process Assessment Scales by Group over Time (*N* = 26)

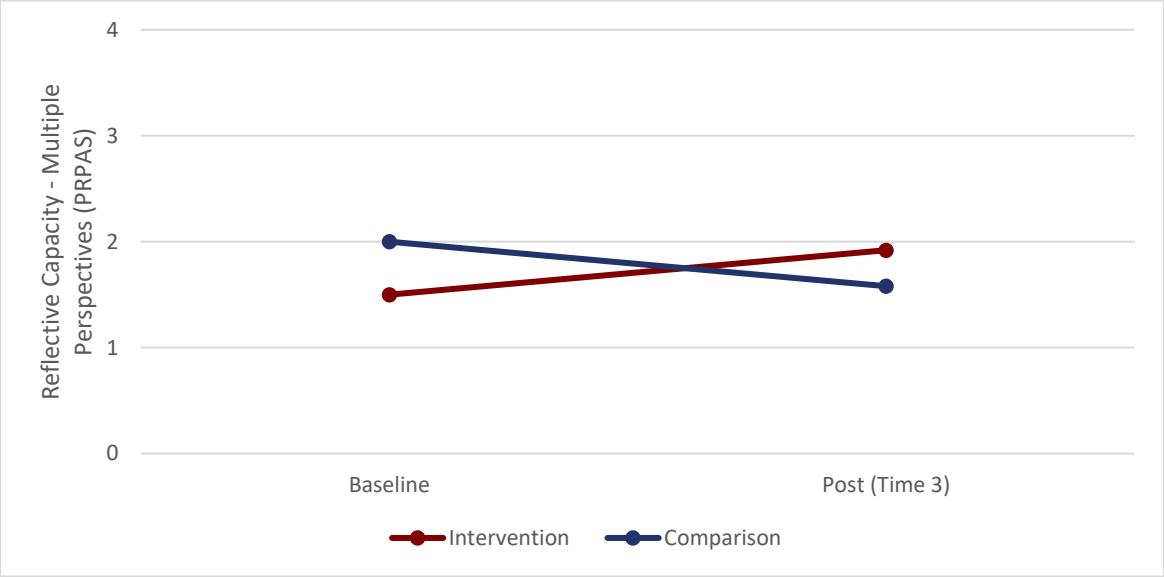
PRPAS Scale and Score Range	Group	Baseline (Time 1) Mean (*SD*)	One-Year Follow-up (Time 3) Mean (*SD*)	F	*p*	η_p^2
Self-Knowledge (Range: 0–8)	Intervention	3.57 (0.93)	3.92 (0.86)	1.28	.269	.05
	Comparison	3.69 (1.11)	3.46 (1.54)			
Self-Regulation Range: 0–12)	Intervention	4.42 (1.13)	5.46 (1.03)	2.86	.104	.11
	Comparison	5.07 (1.15)	4.92 (1.44)			
Multiple Perspectives (Range: 0–4)	Intervention	1.50 (0.67)	1.92 (0.57)	7.28	.013*	.23
	Comparison	2.00 (0.67)	1.58 (0.49)			
Collaboration (Range: 0–12)	Intervention	4.00 (1.64)	4.81 (1.59)	1.49	.235	.06
	Comparison	4.65 (1.64)	4.46 (1.09)			
Process (Range: 0–8)	Intervention	3.00 (0.84)	3.69 (0.90)	1.91	.180	.07
	Comparison	3.58 (1.13)	3.46 (1.38)			
Authentic Attitude (Range: 0–12)	Intervention	3.12 (1.78)	4.46 (1.20)	2.59	.121	.10
	Comparison	4.34 (1.07)	4.46 (0.95)			

Note. The sample for the PRPAS was the focused sample of staff: Intervention (*N* = 13) and Comparison (*N* = 13). The response scale for the individual items that make up each of the PRPAS scales ranges from 0 to 4 with higher cores being more desirable. The items in each scale are summed, resulting in possible scores ranging from 0 to 4 (1-item scale), 0 to 8 (2 item-scale), or 0 to 12 (3-item scale), depending on the number of items in each scale.

Most of the scores on the PRPAS fell around the middle or a little below the possible range of scores, indicating some reflective capacity took place but that there was still room for growth. Overall, scores increased between the two time points for the intervention group and decreased or stayed the same for the comparison group. However, only the Multiple Perspectives scale showed a significant interaction effect of group by time, showing an increase between the two time points for staff in the intervention group versus staff in the comparison group. This scale assesses "the extent to which the respondent is aware of the personal history, experiences, and culture of self and strives to understand those of the client and other important people in client's life and to help the client understand these differing perspectives and their impact on behavior (Heller, 2017)." Thus, staff receiving the intervention increased in their ability to see a

situation from multiple perspectives over time, on average, while staff in the comparison group did not significantly change over time on this measure (see Figure 8). The sample was small, so large effects were necessary to yield sufficient power to reach significance.

Figure 8. Reflective Capacity - Multiple Perspectives by Group over Time (*N* = 26)



Note. The sample for the PRPAS was the focused sample of staff, Intervention (*N* = 13) and Comparison (*N* = 13). The response scale for the single-item Multiple Perspectives scale is 0–4, with higher scores being more desirable. Repeated measures ANOVA found significant group differences over time, $F = 7.28$, $p = .013$.

Examples of Changes in Reflective Capacity

The qualitative interviews with staff, supervisors, and consultants provided numerous examples of changes in reflective capacity over time, which support these quantitative results. It took time for staff to become reflective and to be reflective at work. Some staff were initially resistant to the idea of reflective practice. For example, a teacher said she was not interested in being reflective, she wanted a consultant who provided answers about how to work with specific children. Some staff were less resistant but found the concept of being reflective difficult to understand or apply in their work.

Yet, overall, the qualitative data showed that at most intervention programs, a shift occurred over the course of implementation toward increasing comfort with and increasing capacity for reflective practice. Components of reflective practice that emerged from the interview data include: 1) greater ability to explore issues, engage in active listening, and ask questions; 2) greater ability to think critically about one's reactions and consider one's biases; 3) greater ability to consider others' perspectives; and 4) greater ability to establish or improve boundaries and be mindful of self-care. We briefly describe examples for each of these aspects below:

Active listening and deeper exploration. Staff at Time 3 reported they listen and explore issues more deeply at most of the intervention sites, often spurred by questions asked by the consultants. Some staff remarked upon their consultant's ability to get them "to open up in their team meetings." One staff member shared that, through her work with the consultant, she was able to analyze and delve into her frustrations, breaking them down into parts. At a different site, a staff member noted, "The way she [consultant] talks to me, the things she does ask me and the way she asks me actually does get me to open up more or explore a little further than anyone else ever has." Some interviewees spoke about how they applied these skills directly to their work with children and families. An experienced home visitor realized that in the past year she became more "self-aware" and more "conscious of things that [families] may not point out."

Thinking critically about oneself. Another theme in the staff interviews was greater self-awareness, introspection, and critical thinking about their reactions in the context of their work. For example, a home visitor commented that her consultant asks questions that "help me really think about it more myself. She spends more time on making me think about myself, how I feel about something." In another example, a teacher said that for her the biggest benefit of consultation was "forcing me to pinpoint my challenges, forcing me to pinpoint how I need to work on them, and really articulating that."

Considering the perspectives of others. Another aspect of reflective practice that emerged in the qualitative data was the growing awareness of and attunement to other people's perspectives. Many staff spoke about being more open to considering the struggles of parents and then being able to respond to them differently. At a home visiting program, a staff member praised the consultant for reminding the staff to think about how a parent feels and why the parent might be resistant or act in a certain manner. One home visitor commented that a trauma training by the consultant helped her become less judgmental and more empathetic. She realized that just because she might find something easy, her clients might not view it in the same way; she needed to think about the situation from their perspective.

Similarly, teachers reported coming to view children with challenging behaviors from a different perspective. For example, a teacher noted that she felt more tolerant of things she could not change. The teacher tried to broaden her understanding of the families in the community and the reasons why some children may behave in certain ways. Another teacher spoke about her increased awareness of familial stress and noted that she was slower to make referrals than she had been, allowing children time to grow on their own.

Managing boundaries and self-care. Teachers and home visitors often reported that consultation helped them set or improve boundaries in their work, be it with other staff or with families. At one program, a teacher described learning the importance of boundaries so she was not always feeling stressed and overworked. She had been in the practice of responding to messages on

the communication app used by the school and parents immediately, even during school breaks. Over time she felt less need to respond quickly to every message. Likewise, after a consultant helped her establish boundaries with a family, a home visitor shared that she became calmer at visits and more comfortable with not having all of the solutions to the family's problems.

Staff Burnout

We measured burnout with the Maslach Burnout Inventory (MBI; Maslach et al., 1996), which includes subscales for emotional exhaustion (feeling emotionally overextended by one's work), depersonalization (unfeeling and impersonal toward others), and personal accomplishment (feelings of competence and achievement in one's work; Maslach et al., 1996). Internal consistency was very high for Emotional Exhaustion (Cronbach's alpha = .93, .92, .93, .94 at baseline, T2, T3, T4, respectively), adequate for Depersonalization (Cronbach's alpha = .72, .70, .71, .79 at baseline, T2, T3, T4, respectively), and good for Personal Accomplishment (Cronbach's alpha = .76, .82, .85, .77 at baseline, T2, T3, T4, respectively). Results of the LMM analyses are presented in Table 8.

Table 8. Burnout: LMM Analysis Results (*N* = 107)

Scale	Intervention effect (Time*Group)	Effect of any other variables
Maslach Burnout Inventory – Emotional Exhaustion	β = 0.13, *p* = .961	Staff race (White), β = 6.94, *p* = .048* Staff education, β = 4.15, *p* = .022*
Maslach Burnout Inventory – Depersonalization	β = -0.49, *p* = .639	N/A
Maslach Burnout Inventory – Personal Accomplishment	N/A	N/A

*$p < .05$

Note. Staff scores on the Emotional Exhaustion subscale (*M* = 17.33) tended to be similar to scores in another study with early care and education providers (Carson et al., 2010; *M* = 15.30), slightly lower than a study with preschool teachers, (Jennings, 2015; *M* = 23.83), and slightly higher than a study with home visitors (Spielberger et al., 2019; *M* = 13.2). Range 0–54.
Staff scores on the Depersonalization subscale (*M* = 4.87) tended to be slightly higher than scores of preschool teachers in Jennings (2015; *M* = 3.80) and home visitors in Spielberger et al. (2019; *M* = 1.57). Range 0–30. Staff scores on the Personal Accomplishment subscale (*M* = 35.88) tended to be slightly lower than other studies of preschool teachers (Jennings, 2015; *M* = 40.55) and home visitors (Spielberger et al., 2019; *M* = 40.83). Range 0–48.

All of the scores on the three subscales reflected fairly low burnout levels. Moreover, the intervention did not have a significant effect on any of the three burnout subscales. As shown in Table C-1, staff in both the intervention and the comparison group scored fairly low on the Emotional Exhaustion and Depersonalization subscales and fairly high on the Personal Accomplishment subscale. There were no significant differences between the two groups on the three subscales.

However, other variables did affect some of the subscales. There was an effect of race and ethnicity on the Emotional Exhaustion subscale, after controlling for the effects of the other variables in the model. In particular, staff who identified themselves as White had a higher average score on the Emotional Exhaustion subscale compared to all other racial groups. This means that staff who identify themselves as White reported experiencing higher burnout in the form of Emotional Exhaustion than staff who identify as Black or Hispanic.

There was a positive effect of educational attainment on the Emotional Exhaustion subscale. Staff members with at least a bachelor's degree had a higher average score than those with less than a bachelor's degree. Thus, staff members with a higher level of education reported experiencing greater emotional exhaustion than those with lower levels of education. No significant effects were found on the Depersonalization subscale. (The model did not run for the Personal Accomplishment subscale because staff in both the intervention and comparison groups had very similar means on this subscale at all four time points.[25])

Reflective Functioning and Burnout

The theory of change suggests that there might be a relationship between reflective functioning and well-being, and, specifically, burnout. Thus, we also examined whether there was a relationship between reflective functioning and the three burnout subscales; that is, does an increase in reflective functioning predict lower levels of burnout? First, we ran correlational analyses between the change in RFQ Certainty and Uncertainty subscales between baseline and Time 4 and the MBI subscales at Time 4. None of the correlations were significant (see Table C-6).

Next, we tested the other measure of reflective functioning, the PRPAS, analyzing the correlations between PRPAS change scores and burnout. Because we have PRPAS scores at baseline and Time 3, we were interested in whether improvement in PRPAS scores from baseline to Time 3 was associated with MBI scores at Time 3 (see Table C-7). Staff change in the Process scale (PRPAS, baseline to Time 3 change score) was negatively correlated with the Emotional Exhaustion domain (MBI) at Time 3 (r = -.49, p = .018, n = 23). In addition, staff change in the Collaboration scale (PRPAS, baseline to Time 3 change score) was negatively correlated with the Emotional Exhaustion domain (MBI) at Time 3 (r = -.48, p = .019, n = 23). (The PRPAS was part of the interview, thus this measure was administered only to the focused sample of staff.)

To better understand the impact of the intervention and reflective capacity on burnout for the focused sample, we ran hierarchical multiple regression analyses testing the two PRPAS scales

[25] The final model for Personal Accomplishment could not be computed because there was no convergence using 10 iterations. This could be because there was no effect of time on the average score of the Personal Accomplishment subscale, as shown in the consistent mean scores over time on this subscale in Appendix C, Table C-1.

with change scores that were significantly correlated with the MBI's Emotional Exhaustion domain at Time 3: Process and Collaboration scales. In the regression models predicting Emotional Exhaustion at Time 3, Intervention group was entered into the model in the first step and contributed significantly to the regression models ($F(1, 21) = 7.38$, $p = .013$).

In step two, increase in reflective functioning (PRPAS scales Process and Collaboration, baseline to Time 3 change scores) contributed significantly to the regression models (i.e., change in R^2 was significant when the PRPAS scale was added to the model). Participant educational attainment was added in step three, as education was found to correlate to burnout. Adding education did not significantly change the R^2 and did not contribute significantly to the models. These models demonstrate that improvement in reflective capacity from baseline to Time 3 contributed to lower levels of emotional exhaustion, yet the intervention was the strongest predictor of lower emotional exhaustion at Time 3.

Table 9 presents the hierarchical multiple regression model in which the intervention and the change in reflective process predicted emotional exhaustion at Time 3 ($F(3,19) = 4.50$, $p = .015$). Table 10 presents the model in which the intervention and the change in reflective collaboration predicted emotional exhaustion ($F(3,19) = 4.81$, $p = .012$).

Table 9. Intervention and Change in Reflective Process Predicting Burnout (*N* = 23)

		Δ R^2	Final model β
Emotional exhaustion, T3			
Step 1	Intervention	.260*	-.420*
Step 2	Process change (T1-T3)	.134*	-.336^
Step 3	Education	.022	.154
	Total R^2 = .415*		
	(adjusted .323)		

*$p < .05$, ^$p < .10$

Note. The sample for this model was the focused sample of staff, as the Process variable is a scale in the PRPAS, administered as part of the staff interview.

Table 10. Intervention and Change in Reflective Collaboration Predicting Burnout (*N* = 23)

		Δ R^2	Final model β
Emotional exhaustion, T3			
Step 1	Intervention	.260*	-.427*
Step 2	Collaboration change (T1-T3)	.138*	-.351^
Step 3	Education	.034	.186
	Total R^2 = .432*		
	(adjusted .342)		

*$p < .05$, ^$p < .10$

Note. The sample for this model was the focused sample of staff, as the Collaboration variable is a scale in the PRPAS, administered as part of the staff interview.

Depression

Staff depressive symptoms were measured by the Patient Health Questionnaire, two-item version (PHQ-2; Kroenke et al., 2003). Internal consistency was adequate to good on this scale (Cronbach's alpha = .84, .72, .78, .75 at baseline, T2, T3, T4, respectively). The effect of the intervention on the PHQ-2 scale was not significant (see Table 11). The only variable that showed some effect on depression was teacher role. In particular, after controlling for the effects of the other variables in the model, there was a trend of lead teachers having a lower average score on the PHQ-2 than teacher assistants and home visitors. This means that lead teachers tended to report experiencing depressed mood less frequently than teacher assistants and home visitors.

Table 11. Depression: LMM Analysis Results (*N* = 107)

Scale	Intervention effect (Time*Group)	Effect of any other variables
Patient Health Questionnaire-2	β = 0.037, *p* = .921	Staff role (lead teacher), β = -0.51, *p* = .090^

^*p* < .10

Note. The Patient Health Questionnaire-2 has possible scores ranging from 0 to 6 with higher scores (3+) suggesting need for further screening for depression. Staff scores on the Patient Health Questionnaire tended to be low (*M* = 0.90). Although no published studies have administered the PHQ-2 to early childhood teachers or home visitors, the sample in this study had similar scores as parents of infants in Katch & Burkhardt (2021; *M* = 0.73) and lower scores than pregnant women in Smith et al. (2010; *M* = 2.18).

Self-efficacy and Knowledge

The staff survey included two measures of self-efficacy: The Teacher Opinion Scale (TOS; Geller & Lynch, 1999) and the Goal Achievement Scale (GAS; Alkon et al., 2003), which has been used in previous studies of IECMHC (for example, Egeren, et al. 2011). The TOS measures early childhood providers' feelings of confidence in managing challenging behaviors and their ability to make a positive difference in the lives of children. The GAS measures staff sense of competence in their role as teachers and the behaviors related to the teachers' ability to manage child behavior. Internal consistency was low for the TOS (Cronbach's alpha = .64, .72, .64, .63 at baseline, T2, T3, T4, respectively), indicating that the TOS does not successfully measure one dimension or construct (i.e., self-efficacy). Shivers (2011) found two factors in the TOS with a principal component analysis, yet our data did not align with those two factors. Internal consistency was high for the GAS (Cronbach's alpha = .83, .86, .86, .84 at baseline, T2, T3, T4, respectively). Analysis of the TOS and GAS excluded the sample of home visitors, as many items

on the scales did not apply to them and the final sample of home visitors was small.[26] Results of the LMM analysis are presented in Table 12.

Table 12. Teacher Self-Efficacy: LMM Analysis Results (*N* = 95)

Scale	Intervention effect (Time*Group)	Effect of any other variables
Teacher Opinion Survey (TOS)	β = 0.18, *p* = .809	Staff age, β = -9.67, *p* = .050^
Goal Achievement Scale (GAS)	N/A	N/A

^*p* < .10

Note. Teacher (*M* = 20.94) and home visitor (*M* = 21.86) scores on the GAS tended to be slightly lower than scores in other studies, (e.g., Egeren et al., 2011; *M* = 23.30). The range of GAS scores was 0 to 26. Teacher scores (*M* = 46.91) and home visitor scores (*M* = 45.06) on the TOS tended to be similar to scores in Egeren et al. (2011; *M* = 47.10). The range of TOS scores was 12 to 60.

There was very little difference in TOS and GAS scores between the intervention and comparison groups and little variation in the scores over time. As shown in Table C-1, scores for both groups were around 46 or 47 at each time point on the TOS, which has a possible score range between 12 and 60. On the GAS, which has a possible score range between 0 and 26, teachers in both the intervention and comparison groups scored around 20 or 21 at each time point.

On the TOS, therefore, the LMM analysis showed no effect of the intervention on teachers' self-efficacy. There was a trend toward a negative effect of teacher age on the TOS after controlling for the effects of the other variables in the model. In particular, older teachers had lower average scores on the TOS compared to younger teachers, suggesting that older teachers tended to feel less confident about managing challenging child behaviors than younger teachers.

The LMM analysis of the GAS did not produce results, because the model predicting the GAS could not be computed due to lack of convergence in 10 iterations. However, considering the descriptive statistics (see Table C-1), we did not expect to find group differences given the similar patterns of scores over time for the intervention and comparison groups.

The survey also included a measure of staff's perception of whether they gained knowledge and strategies related to child social-emotional development, which is a key component of IECMHC. The measure, an adapted version of the Social and Emotional Development Inventory (SEDI; Shivers, 2011), was based on self-evaluation tools from the Center on the Social and Emotional

[26] We adapted the TOS and GAS for use with home visitors, changing several of the items on both scales, and included it in the survey administered to home visitors. However, due to the small number of home visitors in the final sample, we could not analyze the home visitors' scores on these scales separately.

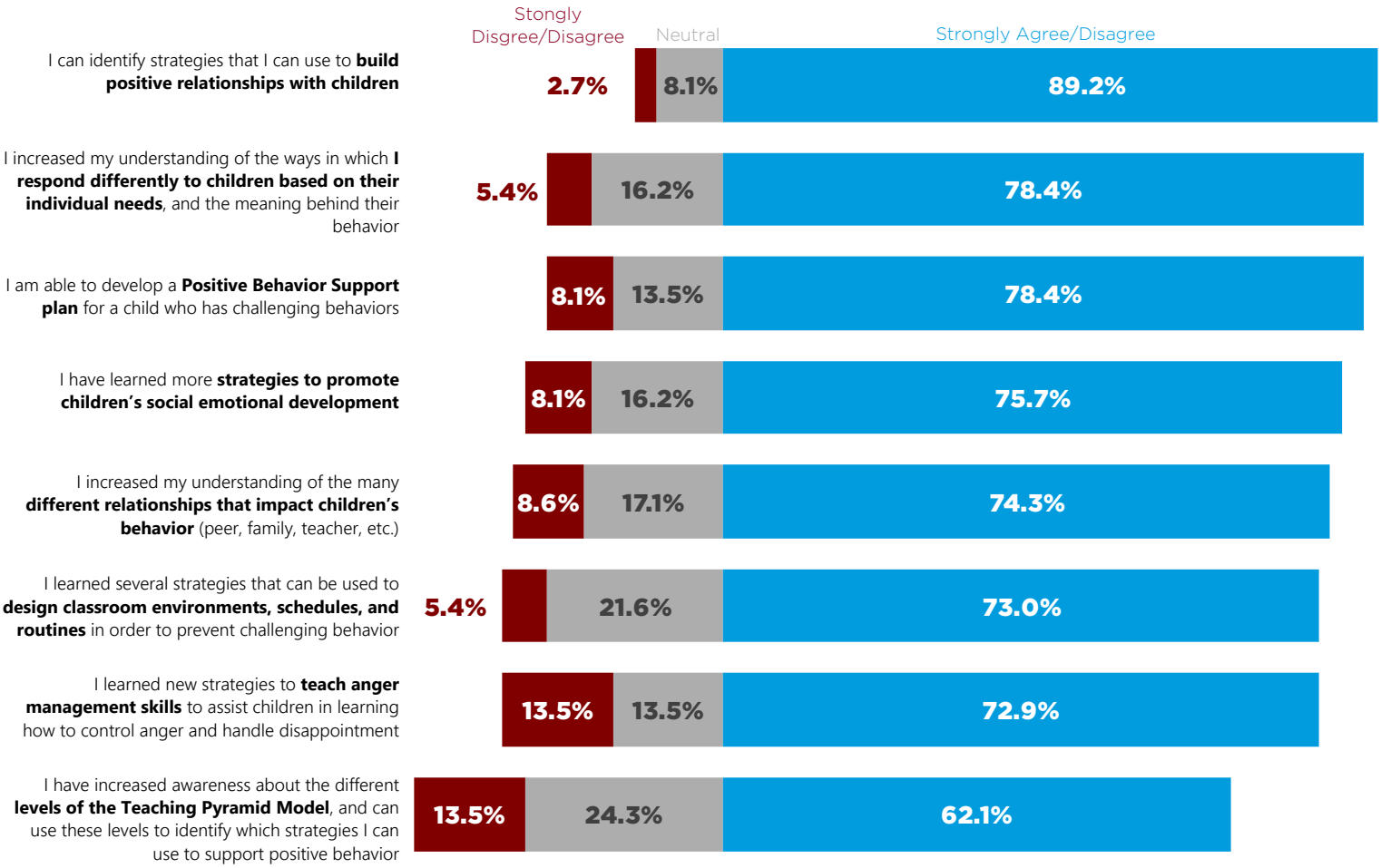
Foundations for Early Learning (CSEFEL).[27] The SEDI items are intended to capture staff's self-assessment of knowledge and skills they may have acquired over the course of the implementation. Staff were only administered these items if they responded that they reported currently receiving mental health consultation.

Table 13 presents responses for staff in the intervention group at Time 4, representing staff self-assessment of knowledge and skills gained over the course of the intervention.[28] As these responses indicate, a majority (62% to 89%) of the staff receiving the Illinois Model of IECMHC agreed that they had increased their understanding of children's social and emotional development and learned new strategies for working with young children. The item that elicited the least amount of agreement among the respondents was an item pertaining to their awareness of the different levels of the Teaching Pyramid Model and ability to use these levels to identify strategies to support positive behavior.

Qualitative data indicated that some staff in the intervention group felt more confident in their abilities by the end of the evaluation, suggesting a potential boost in self-efficacy from the IECMHC. For example, during the final interview, a supervisor explained that the teachers were now more comfortable communicating with parents without the consultant present. The teachers took the strategies and ideas that they learned from the consultant to meetings with parents. The supervisor emphasized this change in the teachers saying, "Before they were like, 'No, we need help. We need help.' Now they are like, 'No, let us try first; give us a chance.'"

[27] CSEFEL is a national resource center funded by the Office of Head Start and Child Care Bureau focused on promoting the social emotional development and school readiness of young children birth to age 5.
[28] Although this measure was administered to staff in the comparison group who responded that they did have access to a mental health consultant, we do not include their responses here because we are following an intent-to-treat approach in the analysis. We also have limited data on the nature of the consultation received by the comparison group, and it would be difficult to interpret any differences or lack of differences we obtained.

Table 13. Social and Emotional Development Inventory (SEDI): Intervention Group Self-Assessment at Time 4 (*N* = 37)



Staff Supervision and Relationships with Supervisors

Encouraging positive relationships between supervisors and staff and reflective supervision are important components of the Illinois Model, as it can improve staff mental health (Susman-Stillman et al., 2020) and the quality of services provided to families (Heffron, 2005). Below we describe the nature of staff supervision at the study programs, as reported by staff, and then the results of measures of reflective supervision and staff–supervisor relationships.

Staff Reports of Supervision Received

The online survey asked staff about the format and frequency of their supervision and its adequacy to meet their needs in four areas of their work. At baseline, the intervention and comparison groups were similar in terms of their experiences with supervision. Over two-thirds of staff in both groups reported receiving supervision on some regular basis, typically once a

month, for 30 to 60 minutes per session. About a third of the staff receiving supervision received group supervision and about a quarter or less received individual supervision. The other staff who received supervision described it as irregular, unscheduled, or as needed. As discussed later in the chapter, supervisors also reported a mix of formats for providing supervision to staff. Because not all staff and supervisors in the pilot study responded to the survey, we cannot compare the responses of the two groups. Importantly, of all the survey respondents, a small percentage reported receiving regular one-on-one supervision.

At baseline, regardless of format, more than two-thirds of those receiving supervision said that these sessions addressed program/administrative issues and professional development needs "very well" or "fairly well." About the same proportion said supervision helped them process their feelings and reactions to their work with children and families. A smaller proportion—just over half of the staff in each group—responded that their supervision addressed clinical issues such as working with children with challenging behaviors "very well" or "fairly well."

Change in Supervision over Time

Structure and frequency of supervision. Over the 15-month implementation period, the percentage of staff receiving any form of supervision remained fairly constant. During the period, about a third of staff responding to the survey said that they did not receive any supervision. However, the staff in the intervention group who did receive supervision reported changes in the structure and frequency of their supervision over time. Specifically, the percentage of staff reporting that they received individual supervision grew from 31% at baseline to 60% at Time 4 (see Table 14).

A chi-square analysis of the Time 4 responses to the question about the primary way staff receive supervision showed a significant difference by group (χ^2 = 11.183, *df* = 3, *p* = .011) That is, staff in the intervention group were much more likely to receive regular one-on-one supervision than staff in the comparison group.

Table 14. Primary Way of Receiving Supervision over Time

	Intervention				Comparison			
Supervision Format	Baseline (*n* =49)	Time 2 (*n* = 49)	Time 3 (*n* = 39)	Time 4 (*n* = 25)	Baseline (*n* = 45)	Time 2 (*n* = 39)	Time 3 (*n* = 35)	Time 4 (*n* = 29)
Regular one-on-one supervision (%)	31	47	64	60	22	23	26	17
Regular group supervision (%)	37	29	23	20	36	28	40	35
Unscheduled supervision (%)	32	24	13	20	40	47	34	41

Adequacy of Supervision. Staff's ratings of the adequacy of their supervision to address various aspects of supervision were similar over time. Table 15 indicates that, on average, staff regarded their supervision in four areas as between "somewhat adequate" and "fairly adequate." At the same time, the response range and standard deviations indicate that there was considerable variability within the ratings at each time point.

Table 15. Adequacy of Supervision over Time

	Intervention				Comparison			
Adequacy of Supervision to address:	Baseline (*n* = 49)	Time 2 (*n* = 49)	Time 3 (*n* = 39)	Time 4 (*n* = 25)	Baseline (*n* = 45)	Time 2 (*n* = 39)	Time 3 (*n* = 35)	Time 4 (*n* = 29)
Professional development								
Mean (*SD*)	3.1 (1.02)	2.7 (1.06)	3.1 (1.05)	3.2 (0.97)	3.2 (0.89)	3.1 (0.88)	3.1 (0.9)	2.9 (0.92)
Range	1–4	0–4	0–4	1–4	0–4	1–4	1–4	1–4
Administrative issues								
Mean (*SD*)	2.7 (1.15)	2.6 (1.21)	3.1 (1)	2.7 (1.14)	2.8 (0.9)	2.9 (0.93)	2.9 (0.96)	2.6 (0.99)
Range	0–4	0–4	0–4	0–4	0–4	1–4	1–4	0–4
Clinical issues								
Mean (*SD*)	2.6 (1.13)	2.5 (1.31)	3.0 (1.07)	2.9 (1.12)	2.7 (1.05)	2.7 (0.98)	2.7 (1.24)	2.5 (1.06)
Range	0–4	0–4	0–4	1–4	0–4	0–4	0–4	0–4
Feelings/reactions to work								
Mean (*SD*)	2.8 (1.12)	2.7 (1.21)	3.1 (1.07)	3.2 (1.07)	2.9 (1.08)	2.9 (1.03)	2.9 (1.08)	2.8 (0.99)
Range	0–4	0–4	0–4	0–4	0–4	0–4	1–4	0–4

Note. Response scale: 0, "Not at all adequate"; 1, "A little adequate"; 2, "Somewhat adequate"; 3, "Fairly adequate"; and 4, "Very adequate."

Reflective Supervision

The supervisory relationship was assessed with two measures. The Reflective Supervision Rating Scale (RSRS; Ash, 2010) was completed by staff, while the Supervisory Worker Alliance Inventory (SWAI; Efstation et al., 1990) was completed by staff and supervisors. (We present the supervisor data in the next section of this chapter.) The SWAI contains subscales measuring the relationship between the staff and supervisor (Rapport) and the extent to which the staff perceives the supervisor encourages focused efforts toward specific goals and tasks expected to benefit clients (Client focus). See Table 16.

Table 16. Staff Views of Supervisory Relationships: LMM Analysis Results (*N* = 107)

Scale	Intervention effect (Time*Group)	Effect of any other variables
Reflective Supervision Rating Scale	β = -0.77, *p* = .751	Staff role (lead teacher)*Group, β = -7.27, *p* =.025*
Supervisory Worker Alliance Inventory – Rapport	β = 0.24, *p* = .573	Center*Time, β = 0.73, *p* = .040* Program size*Time, β = -0.85, *p* = .016* Staff race (White)*Time, β = -0.36, *p* = .047* Staff age*Time[5], β = -0.31, *p* = .006** Staff role (lead teacher)*Group, β = -0.82, *p* = .099^
Supervisory Worker Alliance Inventory – Client Focus	β = -0.23, *p* = .567	Staff role (lead teacher)*Group, β = -0.98, *p* = .055^ Staff age*Time[5], *β* = -0.22, *p* = .073^

^*p* < .10, **p* < .05, ***p* < .01

Note. Staff scores on the RSRS (*M* = 41.38) tended to be slightly lower than Spielberger et al.'s (2019) study of home visitors (*M* = 44.75). Range 17–51.

Staff scores on the Rapport subscale of the SWAI (*M* = 5.54) were similar to psychology interns in Efstation et al. (1990; *M* = 5.85). Range 1–7.

Staff scores on the Client Focus subscale of the SWAI (*M* = 5.39) were similar to psychology interns in Efstation et al. (1990; *M* = 5.44). Range 1–7.

Reflective Supervision Rating Scale (RSRS)

The intervention did not have a significant effect on the RSRS scores. At baseline, both the intervention and the comparison groups rated their supervision similarly, around 42. Given that the upper range of the RSRS is 51, this means staff responded positively about their supervision at baseline. The data then show very little change in ratings over time by either group. At the same time, the LMM analysis found that teacher role was a significant factor in the RSRS scores. Specifically, there was a negative effect of receiving the intervention on the lead teachers' rating of reflective supervision, after controlling for the effects of the other variables in the model. Lead teachers who received the intervention were predicted to have an average score on the RSRS that is 7.3 units *lower* than lead teachers in the comparison group. This means that lead teachers who received the intervention perceived the quality of the reflective supervision they received to be lower than lead teachers in the comparison group.

Supervisor Working Alliance Inventory (SWAI): Rapport

As shown in Table C-1, scores on the Rapport subscale can range from 1 to 7. Scores for both the intervention and comparison groups were similar and fairly positive over time—between 5 and 6, on average, at each time point. Thus, the effect of the intervention was not significant on the SWAI Rapport subscale. However, a number of other variables had a significant effect on this scale. We found significant two-way interactions between time and the following variables:

program type, program size, race (White), and staff age. The two-way interaction between the intervention and the staff in the lead teacher role was also significant. We describe each of these relationships below.

Time and program type: There was a positive effect of time on the Rapport subscale for the staff working in center-based programs, after controlling for the effects of the other variables in the model. Staff members working in center-based programs had higher scores on the Rapport subscale (on average, over time) compared with staff working in home visiting programs. This means that staff members working in center-based programs perceived supervisors' efforts to build a bond or relationship with them over time to be greater than home visitors' perceptions of their supervisors.

Time and program size: Size of the program interacted with time to effect on the Rapport subscale. There was a negative effect of time on the Rapport subscale for the staff working in larger programs, after controlling for the effects of the other variables in the model. Staff members working in larger programs had a lower average score over time on the Rapport subscale than those working in smaller programs, meaning that staff in larger programs perceived supervisors' efforts to build a bond or relationship with them over time to be poorer than those in smaller programs.

Race: Staff members who are White had an average score on the Rapport subscale that was lower over time than staff who are Black or Hispanic, after controlling for the effects of the other variables in the model. Thus, staff who are White perceived supervisors' efforts to build a bond or relationship with them over time to be lower than staff who are Black or Hispanic.

Lead teacher status: There was a trend toward a negative effect of receiving the intervention on the lead teachers' Rapport subscale, after controlling for the effects of the other variables in the model. In particular, lead teachers who received the intervention trended toward having a lower average score on the Rapport subscale compared to lead teachers in the comparison group. This means that lead teachers in the intervention group did not perceive supervisors' efforts to build a bond or relationship to be as good as lead teachers in the comparison group.

Time and staff age: The effect of time on the Rapport subscale for older staff members appeared to be nonlinear. In particular, according to the final model, a polynomial of grade 5 is the best to describe this relationship over time. This means that the shape of the curve changed direction, increasing and decreasing five times.

SWAI: Client Focus

The effect of the intervention was not significant on the SWAI Client Focus subscale. Here, too, responses of staff in both the intervention and comparison groups were fairly positive at each time point—again, between 5 and 6 on a 7-point scale (see Table C-1). However, other variables

had a significant effect on this scale. We found significant two-way interactions between the intervention and the staff role as lead teacher, and between time and staff age.

Lead teacher status: There was a trend toward a negative effect of receiving the intervention on the lead teachers' Client Focus subscale, after controlling for the effects of the other variables in the model. In particular, lead teachers who received the intervention tended to have a lower average score on this subscale compared to lead teachers in the comparison group. This means that there was a trend toward lead teachers in the comparison group believing their supervisors encouraged focused efforts toward specific goals and tasks expected to benefit clients more so than lead teachers in the intervention group.

Time and staff age: The effect of time on the Client Focus subscale for older staff members appeared to be nonlinear, a polynomial of grade 5, like the Rapport subscale. This means that the shape of the function increased and decreased five times.

Supervisor Outcomes

The consultants spent considerable time meeting with program supervisors individually as well as with individual staff or a team of staff. Based on the theory of change, we expected that consultation delivered to supervisors using the Illinois Model of IECMHC would improve the quality of supervision and supervisors' reflective capacity and well-being (burnout and depression). Because only a small sample of supervisors remained in the evaluation over the course of the initiative, we could not analyze their surveys for differences between the intervention and comparison groups over time. However, information from supervisor interviews provide some insight into the supervision they provide and their well-being. In this section, we describe what supervisors reported about their supervisory practices, results of standardized measures of well-being and relationships, and reflective capacity over time.

Supervisors' Reports of Supervision Provided and Received

Just eight (57%) of the 14 supervisors at intervention programs we could follow over time reported at baseline that they provided some kind of supervision to staff; in comparison, four of five supervisors in the comparison group reported providing some form of supervision. Similar to staff's responses about supervision, supervisors reported that they conduct staff supervision in a variety of ways—individual meetings with staff, group meetings, and unscheduled (as needed) supervision.

We also asked supervisors how often they met with their own supervisor for supervision. At baseline, there was a great deal of variability in the responses for both intervention and comparison program supervisors, ranging from "weekly" to "never." In the intervention group, just over half (57%, n = 8) of supervisors reported meeting at least monthly with their supervisor; the other supervisors at intervention programs said they meet with their supervisor quarterly or

less often. At Times 2, 3, and 4, about two-thirds of the supervisors reported only meeting once a year with their supervisors, and a third responded "never." Although the sample was smaller, the patterns in the comparison group were similar to those in the intervention group. These frequencies suggest that many supervisors in the programs in the evaluation do not regularly receive supervision themselves.

Change in Supervision over Time

Frequency and structure of supervision. Over the 15-month implementation period, supervisors who completed the survey continued to report using similar structures for supervision as they did at baseline. However, supervisors' survey responses suggest an increasing effort to provide supervision to their staff during the implementation period. Specifically, the share of supervisors at intervention programs who reported providing supervision increased from 57% at baseline to 75% or greater in the three follow-up surveys (77% at Time 2, 83% at Time 3, and 75% at Time 4). At baseline, 80% of the supervisors in the comparison programs provided supervision, and the proportion was lower in the follow-up surveys (60% at Time 2, 50% at Time 3, and 75% at Time 4). There were no noteworthy changes over time in terms of the structure of supervision.

Qualitative data from the supervisor interviews support the quantitative findings of an increase in supervision of staff over time, while also acknowledging the challenges of making time for supervision. At Time 3, a supervisor of a home visiting program told us that she had recently started doing reflective supervision once a month, after letting it lapse, because she was missing the individual contact with her staff and hearing how they and their families were faring. At Time 3, a supervisor at an early childhood center-based program reported:

> At first it was kind of a struggle, and then I created a plan where it's already scheduled beforehand–-because things happen, right? So sometimes I might be short of staff and have to go cover a classroom or like a last-minute meeting happens and then I have to reschedule. So that's the only thing that was hard was kind of having it as a secure time where we have to get the supervision done and really respecting it as opposed to where I had to reschedule and move things around. That's the biggest struggle, time management.
> —*Supervisor*

Adequacy of Supervision. Supervisor's ratings of the adequacy of their supervision to address various staff over time fell between "somewhat adequate" and "fairly adequate." There was a slight trend towards a higher rating over time in the intervention group, and a trend towards a lower rating in the comparison group. See Table 17.

Table 17. Adequacy of Supervision Provided over Time[a]

	Intervention (*N* = 14)				Comparison (*N* = 5)			
Adequacy of supervision to address:	Baseline (*n* = 8)	Time 2 (*n* = 8)	Time 3 (*n* = 10)	Time 4 (*n* = 6)	Baseline (*n* = 4)	Time 2 (*n* = 3)	Time 3 (*n* = 2)	Time 4 (*n* = 3)
Professional development								
Mean (*SD*)	2.0 (0.53)	3.0 (0.76)	2.9 (0.74)	2.7 (0.82)	3.3 (0.5)	3.7 (0.58)	3.0 (0)	2.7 (0.58)
Range	1–3	2–4	2–4	2–4	3–4	3–4	3–3	2–3
Administrative issues								
Mean (*SD*)	2.4 (0.92)	3.1 (0.64)	2.9 (0.74)	2.8 (0.75)	2.8 (0.5)	2.7 (1.53)	3.0 (1.41)	1.7 (0.58)
Range	1–4	2–4	2–4	2–4	2–3	1–4	2–4	1–2
Clinical issues								
Mean (*SD*)	1.9 (0.64)	3.4 (0.52)	3.1 (0.74)	2.7 (0.82)	3.0 (0.82)	2.7 (1.15)	3.0 (1.41)	3.0 (1)
Range	1–3	3–4	2–4	2–4	2–4	2–4	2–4	2–4
Feelings/reactions to work								
Mean (*SD*)	2.0 (1.07)	3.4 (0.74)	3.0 (0.94)	2.7 (0.82)	3.0 (0)	3.7 (0.58)	3.5 (0.71)	2.7 (1.15)
Range	0–3	2–4	1–4	2–4	3–3	3–4	3–4	2–4

[a] Response scale: 0, "Not at all adequate"; 1, "A little"; 2, "Somewhat"; 3, "Fairly"; and 4, "Very adequate."

Self-reported Change in Supervision over Time

Starting with the Time 2 survey, all supervisors were asked whether supervision provided to staff had changed in the last 6 months. Intervention program supervisors at all three time points noted that their supervision had "change somewhat" or more. Comparison program supervisors were more likely to report that the supervision they provided to staff changed "a little" at all three time points. (See Table 18.)

Table 18. Frequency of Changes Noted in Supervision Provided in Past 6 Months over Time

Degree of Change	Intervention				Comparison			
	Baseline N/A	Time 2 (*n* = 10)	Time 3 (*n* = 10)	Time 4 (*n* = 6)	Baseline N/A	Time 2 (*n* = 3)	Time 3 (*n* = 2)	Time 4 (*n* = 3)
Changed drastically (%)		10	10	0		0	0	0
Changed a lot (%)		20	20	50		0	0	0
Changed somewhat (%)		50	40	33		0	50	33
Changed a little (%)		10	10	17		100	50	67
Not changed at all (%)		10	20	0		0	0	0

To understand the nature of these changes, we looked to the interviews with supervisors and directors. At Time 4, a director of an early childhood center in the intervention group talked about the way she had changed her supervision of the program supervisors in order to model for them how to supervise the frontline staff. "I think for me the growth has been really being more reflective with the supervisors that I have, sort of modeling the reflective supervision pieces so then they could do it with their staff." At another early childhood center, also at Time 4, a program supervisor explained how the mental health consultant helped change her supervision approach:

> It was hard for me at first to make the switch from personal conversation to more professional. The sessions that we had ended up being more personal than professional. But I really feel like [the consultant] has supported me and given me ideas on how to switch that around and bring [the conversation] back to the classroom. That's really helped. . . . And I just make sure we have those reflective practices in place, and it's not all just focused on the demands of the job. —*Supervisor*

Supervisor Reflective Capacity, Relationships, and Well-being

Standardized Measures

Reflective Functioning Questionnaire (RFQ)

For the Certainty subscale of the RFQ, scores for both supervisors in intervention and comparison programs were higher and remained fairly stable over time. This suggests that supervisors in both program types are comfortable with the ambiguity of all people. For the Uncertainty subscale, scores for both supervisors in intervention and comparison programs were

very low over time, indicating strong reflective capacity overall but no significant change over time. Table 19 presents the two subscale scores for the RFQ at four points in time for program supervisors.

Supervisor Worker Alliance Inventory

The Supervisory Working Alliance Inventory (SWAI) asks supervisors to rate their relationship with staff in three domains: Rapport, a measure of how the supervisor perceives her relationship with her staff; Client Focus, a measure of how the supervisor interprets staff's interest in and understanding of their families; and Identification, a measure of the extent to which supervisors believe staff identify with the supervisors' goals and strategies for working with children and families.

Table 19. Supervisor Scores on the Reflective Functioning Questionnaire

	Intervention (*N* = 14)				Comparison (*N* = 5)			
	Baseline (*n* = 14)	Time 2 (*n* = 13)	Time 3 (*n* = 12)	Time 4 (*n* = 8)	Baseline (*n* = 4)	Time 2 (*n* = 5)	Time 3 (*n* = 4)	Time 4 (*n* = 4)
Certainty								
Mean (*SD*)	2.0 (0.90)	2.3 (0.62)	2.1 (0.73)	2.1 (0.83)	2.4 (0.84)	2.4 (0.56)	2.0 (0.63)	1.9 (1.19)
Range	0.17–3.00	1–3	0.67–2.8	0.67–3	1.17–3.00	1.67–3	1.3–2.5	0.2–2.8
Uncertainty								
Mean (*SD*)	0.1 (0.24)	0.1 (0.16)	0.1 (0.16)	0.1 (0.13)	0.3 (0.32)	0.3 (0.41)	0.1 (0.16)	0.2 (0.34)
Range	0–0.67	0–0.33	0–0.33	00.33	0.00-0.60	0-1	0-0.33	0-0.67

Response scale: 1, "Strongly agree" to 7, "Strongly disagree."

As shown below in Table 20, there was little variability in the three SWAI subscales over time for supervisors in intervention and comparison programs. The scores on all three subscales for both groups were high over time, suggesting that supervisors' perspectives of their relationships were positive to begin with and did not change over time.

Table 20. Supervisor Scores for the Supervisory Worker Alliance Inventory

	Intervention (*N* = 14)				Comparison (*N* = 5)			
	Baseline (*n* = 14)	Time 2 (*n* = 10)	Time 3 (*n* = 11)	Time 4 (*n* = 8)	Baseline (*n* = 5)	Time 2 (*n* = 5)	Time 3 (*n* = 4)	Time 4 (*n* =4)
Rapport								
Mean (*SD*)	5.3 (0.83)	5.4 (0.41)	5.3 (0.81)	5.5 (0.97)	5.7 (0.92)	5.7 (1.10)	5.1 (1.52)	5.5 (0.74)
Range	3.57–6.57	4.9–6.1	3.9–6.4	4–7	4.14–6.57	4–6.9	3–6.3	4.4–6
Client Focus								
Mean (*SD*)	4.7 (0.88)	5.3 (0.80)	4.8 (0.80)	4.4 (0.83)	5.3 (1.26)	4.6 (0.91)	4.5 (1.02)	4.5 (0.85)
Range	3.44–6.44	3.8–6.4	3.1–5.9	3.3–5.7	3.22–6.56	3.7–5.9	3-5.3	4-5.8
Identity								
Mean (*SD*)	5.0 (0.90)	5.2 (0.56)	5.2 (0.99)	4.8 (0.96)	5.3 (1.57)	5.2 (0.86)	4.5 (1.46)	4.9 (0.94)
Range	3.29–6.29	4–6	3.3–6.4	3.1–6.4	2.71–6.86	4–6.3	2.6–6	3.7–5.9

Maslach Burnout Inventory

We used the Maslach Burnout Inventory (MBI; Maslach, 1996) to assess whether working with a mental health consultant would impact supervisors' engagement with their work. Table 21 presents the supervisor scores for the three subscales on the MBI.

On the Emotional Exhaustion subscale, the scores for both groups of supervisors were very low. The scores for supervisors in both types of programs did not change substantially over time, with intervention program supervisors' scores ranging from 14.6 to 15.6 and comparison program supervisors' scores ranging from 8.8 to 12.8.

Table 21. Supervisor Scores for the Maslach Burnout Inventory

	Intervention (*N* = 14)				Comparison (*N* = 5)			
	Baseline (*n* = 14)	Time 2 (*n* = 13)	Time 3 (*n* = 12)	Time 4 (*n* = 8)	Baseline (*n* = 5)	Time 2 (*n* = 5)	Time 3 (*n* = 4)	Time 4 (*n* = 4)
Emotional Exhaustion								
Mean (*SD*)	15.4 (10.04)	15.6 (13.80)	14.6 (10.43)	15.4 (12.89)	8.8 (9.50)	9.8 (12.04)	12.8 (9.35)	10.0 (10.68)
Range	0–37	0–42	0–32	2–37	0–25	1.12–31	4–26	4–26
Depersonalization								
Mean (*SD*)	3.5 (3.06)	2.9 (3.17)	3.4 (3.08)	3.0 (3.66)	3.4 (3.85)	4.0 (8.40)	4.8 (7.50)	4.8 (7.09)
Range	0–10	0–11	0–8	0–9	0–8	0–19	1–16	0–15
Personal Accomplishment								
Mean (*SD*)	37.8 (6.74)	36.5 (7.12)	37.8 (5.65)	32.9 (4.29)	42.4 (4.62)	37.6 (6.11)	35.3 (5.44)	38.3 (4.72)
Range	25–48	22–46	30–47	27–40	36–46	32–46	31–43	35–45

The Depersonalization subscale also had very low scores over time for both groups of supervisors. There was very little difference between supervisors at intervention programs and comparison programs.

The Personal Accomplishment subscale had fairly high average scores for supervisors at both intervention and comparison programs. It is interesting that the intervention program supervisor saw a decline in average scores by Time 4, moving from 37.8 at Time 3 down to 32.9 at Time 4, while the comparison program supervisors increased their scores from an average of 35.3 at Time 3 to 38.3 at Time 4.

Depression: Personal Health Questionnaire (PHQ)

Supervisors in both intervention and comparison programs scored fairly low on the two-item PHQ depression screen over time, indicating that supervisors tended to have few depressive symptoms. The range for the PHQ-2 is 0 to 6, with a clinical cutoff score of 3. Supervisors in both groups averaged a score of 1 or lower. shows the scores for the PHQ at four points in time for both groups of supervisors.

Table 22. Supervisor Scores on the Personal Health Questionnaire

	Intervention (*N* = 14)				Comparison (*N* = 5)			
	Baseline *n* = 14	Time 2 *n* = 13	Time 3 *n* = 12	Time 4 *n* = 8	Baseline *n* = 5	Time 2 *n* = 5	Time 3 *n* = 4	Time 4 *n* = 4
Mean (*SD*)	1.1 (1.33)	0.9 (1.38)	0.3 (0.65)	0.9 (0.99)	0.6 (0.89)	0.0 (0.00)	0.5 (1.00)	0.3 (0.50)
Range	0–4	0–4	0–2	0–2	0–2	0–0	0–2	0–1

Note. The possible range for scores on the PHQ-2 is 0 to 6.

Goal Achievement Scale (GAS)

According to Alkon and colleagues (2003), developers of the GAS, this scale measures teachers' and supervisors' competencies on general mental health activities or program goals.

Table 23 shows the GAS scores for supervisors in center-based and home vising programs. The baseline scores for intervention program supervisors at both types of programs were fairly low. These subsamples were very small, so patterns could not be assessed.

Table 23. Supervisor Scores on the Goal Achievement Scale

	Intervention (*N* = 14)				Comparison (*N* = 5)			
	Baseline	Time 2	Time 3	Time 4	Baseline	Time 2	Time 3	Time 4
Center-based	*n* = 7	*n* = 2	*n* = 3	*n* = 4	*n* = 2	*n* = 2	*n* = 0	*n* = 3
Mean (*SD*)	16.9 (3.76)	24.0 (0.00)	25.0 (1.73)	16.3 (6.60)	16.5 (6.36)	21.5 (3.54)	--	17.0 (na)
Range	13–23	24–24	23–26	9–24	12–21	19–24	--	17–17
Home Visiting	*n* = 5	*n* = 4	*n* = 3	*n* = 3	*n* = 2	*n* = 1	*n* = 2	*n* = 1
Mean (*SD*)	19.8 (5.63)	21.5 (5.74)	25.0 (1.41)	20.0 (6.08)	24.0 (1.41)	23.0 (na)	21.0 (na)	22.0 (na)
Range	13–26	14–26	24–26	13–24	23–25	23–23	21–21	22–22

Note. We adapted the GAS with the author's permission to be used with both center-based and home visiting programs by re-wording two items to make them more applicable in home visiting programs.

Growth in Supervisor Reflective Capacity: Qualitative Interviews

Again, given the small sample, it is not surprising that the results presented above indicate no obvious differences in the standardized measures administered to supervisors. Other, nonstandardized, quantitative data and qualitative data from interviews with staff, supervisors, and consultants did indicate growth in supervisors' reflective capacity and relationships with staff.

Consistent with our analysis of the staff interviews, we also saw evidence of increasing reflective capacity in the supervisors we were able to follow over time. Their interviews also included the themes of 1) active listening and deeper exploration of issues, 2) the ability to think critically about one's reactions and biases, 3) the ability to consider others' perspectives, and 4) the ability to establish or improve boundaries and be mindful of self-care. We briefly describe examples of these themes below:

Active listening, questioning, and exploration. There were numerous examples in their interviews of how supervisors applied these techniques in their work with staff. For example, a supervisor noted that she felt that she had improved at posing "more open and reflective questions" to her staff. Another supervisor shared that they began implementing group reflection as part of their monthly staff meetings to encourage peer support and peer feedback. Still other supervisors at

different sites emphasized the importance of listening to staff more in the later interviews than they did at their Baseline interviews. For example, at her Time 3 interview, a supervisor at an early childhood center said, "I am working harder to listen. I'm working harder to reflect on my own. I'm working harder to include people on decisions. And I'm working harder to ask questions before I make a decision." The consultants also noted changes in listening in administrators. For example, one consultant noted a shift in a director's patience and interactions with her staff and attributed this, in part, to the director's improved listening. She felt that the director was taking more time to listen to her staff and to value the process of her staff talking through issues.

Thinking critically about oneself. A supervisor noted that consultation "gives staff the opportunity to reflect on their own implicit biases, their own reasons for reacting a certain way, or how their own personal experiences or how their own personal backgrounds can shape how they see parents or how they see families or how they see children." Likewise, several supervisors commented that the consultation provided an opportunity for them to think critically about themselves. One went so far as to say that the consultation and her ability to be reflective about decisions she made contributed to her growth as a leader.

Other supervisors commented on changing the way they responded to staff to encourage staff to think and reflect more about an issue before reacting. For example, an early childhood supervisor credited consultation with moving her to think about "supporting your teachers where they're at. Not so much giving them the answers, but getting them to think and reflect on how to handle certain situations in their room." Other supervisors echoed this idea of stepping back and encouraging critical thinking and reflection. Critical thinking and reflection can be as much about the supervisor as about their staff. For example, one supervisor said she reflects back on the techniques she learned from her consultant to keep her from getting "worked up in the moment." Rather, she began to take a step back, look at the situation, and think about the situation before acting on it. Supervisors also step back to consider how they work with their staff. For example, one supervisor said, "I would always give answers instead of prompting them to be more reflective. I think that's where I have learned to take a step back. . . . I have to prompt them to problem-solve and to be reflective on their own work instead of me just telling them what to do. I think [consultation] helped me in that aspect.

Considering the perspectives of others. Supervisors reported that as a result of consultation, they were more aware of the perspectives of their staff. Both supervisors and consultants shared that consultation helped supervisors consider more of their staff members' perspectives and that served to increase collaboration. As a supervisor explained at Time 3:

> Being part of the project has really helped me step back and look at the bigger picture and try to ask more questions, really kind of thicken the story and find

> out what's happening or why [staff] might be having these feelings, and how we can support each other and work through that. —*Supervisor*

Although consultants observed growth in this area, they thought that supervisors were still learning how to apply what they understood about staff perspectives in their interactions with them.

Managing boundaries and self-care. Similar to staff, supervisors often reported that their work with their mental health consultants helped them establish boundaries in their work with staff. A supervisor at an early childhood center said she learned that she could not answer every call or text as soon as she receives it. This shift in approach, which she described as "realizing that everything doesn't have to be perfect all the time," led her to find her job more manageable. Setting boundaries is also a self-care strategy. One supervisor reflected that she reminds herself "not to own any of their feelings or emotions—good, bad, or indifferent." She came to recognize that it was not helpful for her to own her staff's feelings. She could be supportive and sympathize, but still maintain boundaries, which helped her reduce her stress level. This notion resonated for other supervisors as well. One talked about no longer taking the staff's stress home with her; another talked about being able to "maintain a more calm persona" when talking with staff about challenges. Indeed, respondents from about half of the interviewed programs shared that the consultation helped them manage their stress better.

Staff Interactions with Children and Families: Classroom Climate and Home Visit Quality

In this section, we report results on analyses that examined whether the intervention affected the classroom climate in center programs and the home visit quality in home programs. We used two observational tools designed for these environments: we used the CHILD (Gilliam & Reyes, 2017) in the center programs and the HOVRS-A+ (Roggman et al., 2010) in the home visiting programs.

Classroom Observations

We used the CHILD (Gilliam & Reyes, 2017) to measure the mental health climate of the classrooms. Scores can range from -2 (undermining child mental health) to +2 (promoting child mental health). The CHILD contains nine domains:

- Transitions (Cronbach's alpha = .91, .86, .55, .93 at baseline, T2, T3, T4, respectively)
- Directions and Rules (Cronbach's alpha = .91, .92, .92, .95 at baseline, T2, T3, T4, respectively)
- Social and Emotional Learning (Cronbach's alpha = .95, .91, .86, .94 at baseline, T2, T3, T4, respectively)
- Staff Awareness (Cronbach's alpha = .87, .93, .85, .91 at baseline, T2, T3, T4, respectively)

- Staff Affect (Cronbach's alpha = .91, .93, .89, .91 at baseline, T2, T3, T4, respectively)
- Staff Cooperation (Cronbach's alpha = .74, .76, .70, .37 at baseline, T2, T3, T4, respectively)
- Staff-Child Interactions (Cronbach's alpha = .93, .94, .92, .92 at baseline, T2, T3, T4, respectively)
- Individualized and Developmentally Appropriate Pedagogy (Cronbach's alpha = .93, .91, .87, .95 at baseline, T2, T3, T4, respectively)
- Child Behaviors (Cronbach's alpha = .87, .87, .88, .83 at baseline, T2, T3, T4, respectively).

Two auxiliary dimensions were added by the measure developers using items from other domains: Equity (Cronbach's alpha = .94, .94, .91, .93 at baseline, T2, T3, T4, respectively) and Warmth (Cronbach's alpha = .96, .95, .94, .94 at baseline, T2, T3, T4, respectively). After examining descriptive statistics of our sample, we removed variables that were not statistically significant or that were highly correlated with other variables. The following variables were tested in the models predicting the classroom observation scores:[29]

- Teacher age (2 variables)
- Teacher race and ethnicity (4 variables)
- Teacher educational attainment (2 variables)
- Program size (small, medium, large)
- Intervention or comparison group
- Time

We tested multiple models for the CHILD domains. See Table 24 for a summary of findings from the final model selected for each CHILD domain.

The effect of the intervention was significant over time for two of the 11 CHILD dimensions—Directions and Rules, and Equity—and trending toward significance in one dimension—Individualized and Developmentally Appropriate Pedagogy. The Directions and Rules domain measures behavior management characterized by setting, modeling, and enforcing clear, consistent, and developmentally appropriate rules of conduct and applying proactive and positive behavior strategies. Teachers who received the Intervention had a higher average score on the Directions and Rules domain over time than teachers who did not receive the Intervention (see Figure 9).

[29] Details on the analytic approach and the variables included in the models can be found in Appendix A

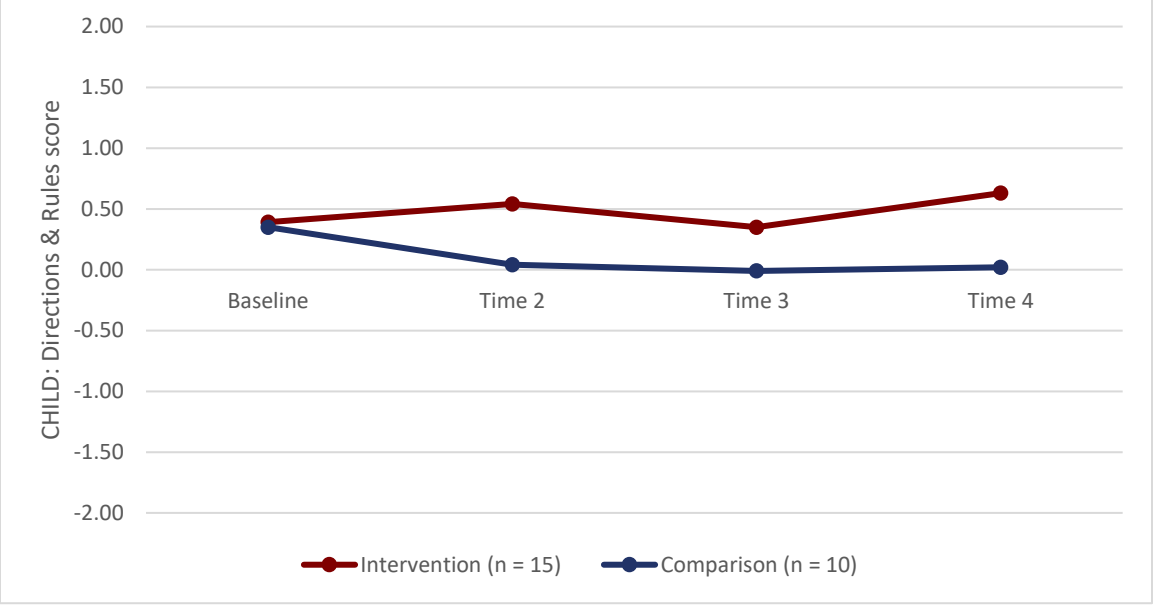
Table 24. Classroom Observations: LMM Analysis Results (*N* = 28 classrooms)

Domain/Auxiliary dimensions	Intervention effect (Time*Group)	Effect of any other variables
Transitions	β = .05, *p* = .824	Teacher race/ethnicity (Hispanic), β = 1.07, *p* = .015* Teacher race/ethnicity (Hispanic) *Time, β = -.49, *p* = .024*
Directions and Rules	β = .22, *p* = .047*	N/A
Social and Emotional Learning	β = .002, *p* = .994	N/A
Staff Awareness	β = .009, *p* = .948	N/A
Staff Affect	β = .13, *p* = .303	Teacher age*Time^3, *p* = .059^
Staff Cooperation	N/A (Model did not run)	N/A
Staff-Child Interactions	β = -.02, *p* = .885	Program size, β = -.28, *p* = .069^ Teacher race (White), β = -.96, *p* = .073^ Teacher age^3, β = -.71, *p* = .040* Teacher age*Time^2, β = -.37, *p* = .077^
Individualized and Developmentally Appropriate Pedagogy	β = .17, *p* = .093^	Teacher education level*Time, β = -.66, *p* = .080^
Child Behaviors	β = -.005, *p* = .982	N/A
Equity	β = .48, *p* = .046* (Main effect)	Teacher race (White), β = -1.03, *p* = .029*
Warmth	β = .077, *p* = .491	Teacher age^3, β = -.82, *p* = .021* Teacher age*Time^2, β = -.35, *p* = .072^

^*p* < .10, **p* < .05

Note. Our usual *p*-level for statistical significance is < .05, but we report results of *p* < .10 to note trends in the data.

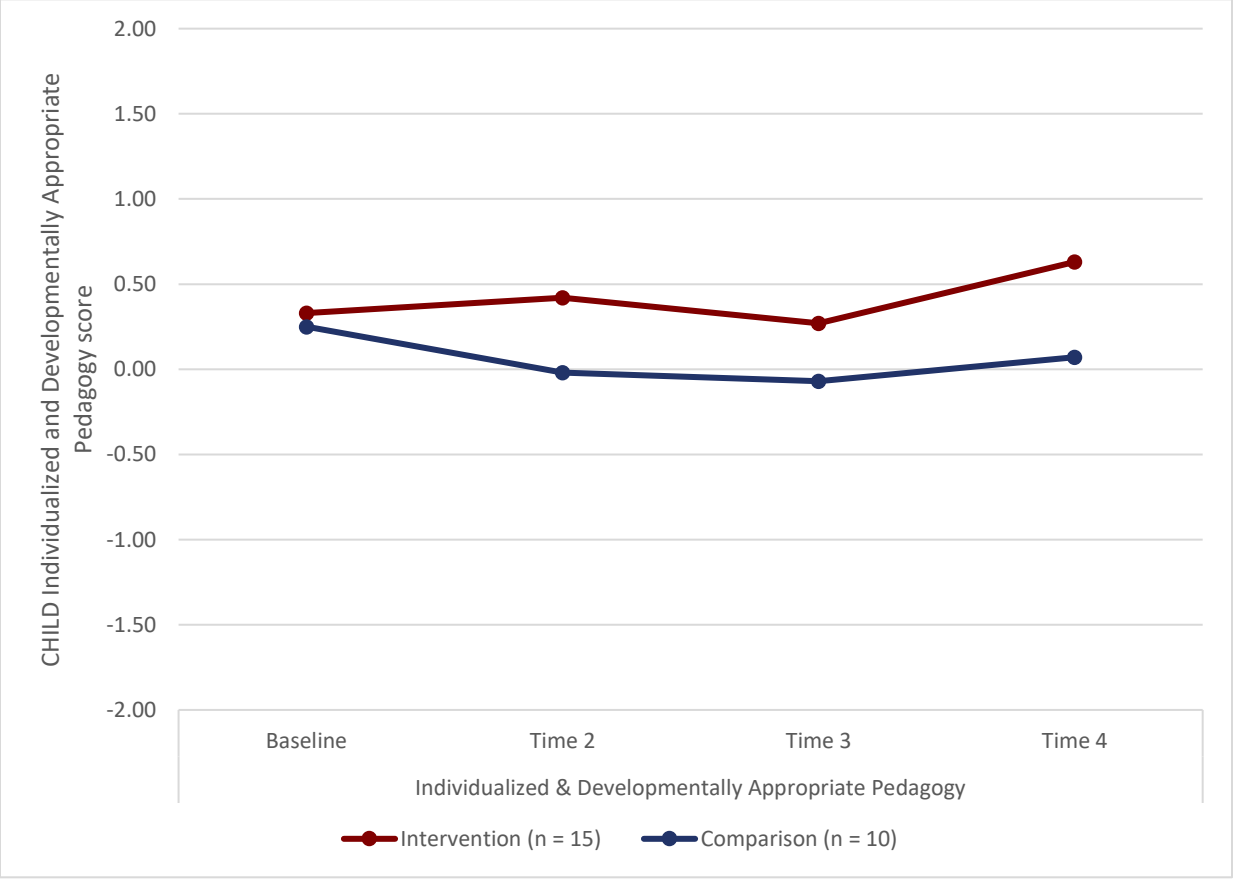
Figure 9. Directions and Rules in Classroom Observations by Group Over Time (*N* = 25)



Note. Scores can range from -2 (undermining child mental health) to +2 (promoting child mental health). Two-way interaction effect was significant between time and group on the Directions and Rules domain, β = .22, *p* = .047*

For the Individualized and Developmentally Appropriate Pedagogy (IDAP) domain of the CHILD, the effect of the intervention was trending toward significance over time. IDAP measures promotion of holistic development through a child-centered and individualized approach. High scores on this domain reflect classrooms in which teachers use strategies that are tailored to each child's needs and activities are centered on children's interests and personal experiences. IDAP scores have been found to moderately correlate with measures of early language and literacy instruction (Reyes & Gilliam, 2018). Teachers who received the intervention tended to have higher average IDAP scores over time than teachers who did not receive the intervention, as displayed in Figure 10. In addition, there was a negative effect of time on the IDAP domain for the teachers with higher educational attainment. Teachers with higher educational attainment had a lower average score on the IDAP domain over time than teachers with lower educational attainment.

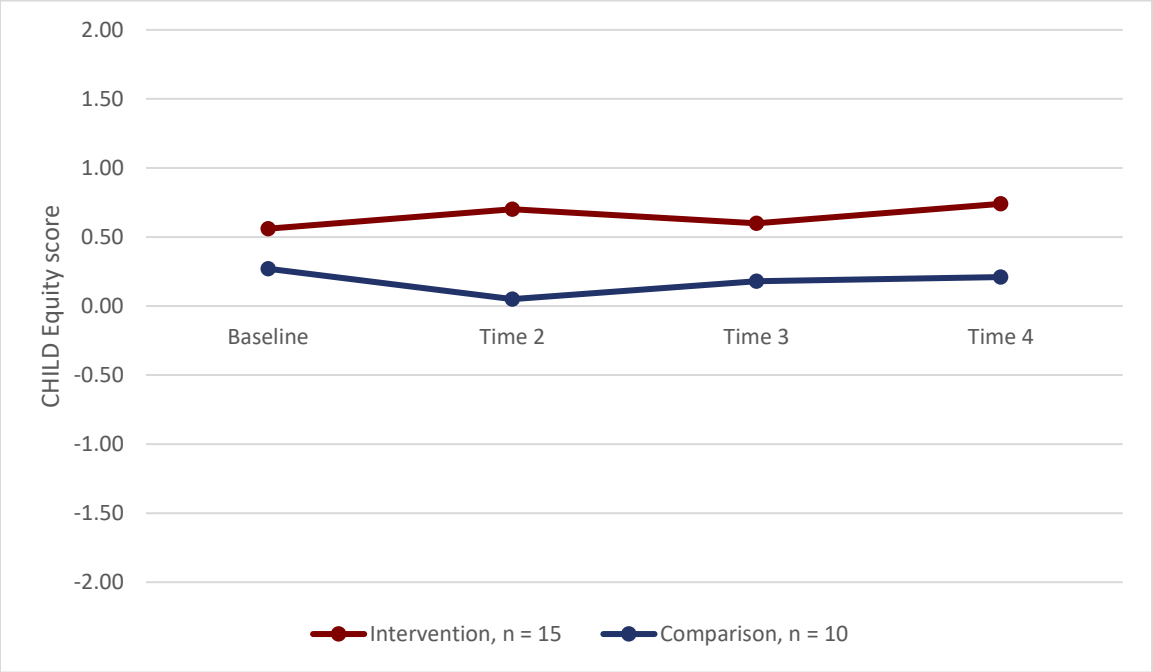
Figure 10. Individualized & Developmentally Appropriate Pedagogy in Classroom Observations by Group Over Time (*N* = 25)



Note. Scores can range from -2 (undermining child mental health) to +2 (promoting child mental health).
Two-way interaction effect between time and group on the Individualized and Developmentally Appropriate Pedagogy domain, β = .17, p = .093.
Our usual *p*-level for statistical significance is < .05. However, we also report results of $p < .10$ to show a trend in the data.

The Equity auxiliary dimension was created by the measure developers and is comprised of items from other domains—two items from Staff Awareness, one item from Staff-Child Interactions, and one item from IDAP.[30] There was a significant effect of the intervention on the Equity auxiliary dimension. Teachers who worked in intervention programs had higher average scores on the Equity auxiliary dimension than teachers in the comparison sites, as seen in Figure 11. In addition, there was a negative effect of race and ethnicity on the Equity domain, after controlling for the effects of the other variables in the model. In particular, teachers who identified as White had a lower average score on the Equity auxiliary dimension than Black and Hispanic teachers.

Figure 11. Equity in Classroom Observations by Group Over Time (*N* = 25)



Note. Scores can range from -2 (undermining child mental health) to +2 (promoting child mental health). Main effect of the intervention on the Equity auxiliary dimension β = .48, *p* = .046.

Additional significant findings in the classroom observations were related to other variables. Hispanic teachers had an average score on the Transitions domain that was higher than non-Hispanic staff members. However, Hispanic teachers' average scores on the Transitions domain decreased over time. Scores on the Staff-Child Interactions scale were affected by program size (larger size was associated with poorer interactions); teacher race and ethnicity (non-Hispanic

[30] Equity auxiliary dimension is comprised of the following items: (1) staff physically circulates around children; (2) staff notices children's overt and subtle signals for assistance; (3) staff attends to children equitably; (4) staff provides appropriate individualized support and feedback.

White teachers had poorer interactions); and staff age. The effect of staff age on Staff-Child Interactions, Staff Affect, and Warmth was nonlinear, meaning that there was not a direct relationship between staff age and these dimensions.

Home Visit Observations

Home visits were video recorded with different families at four times during the study, resulting in a total of 42 videos across seven home visitors (see Table 24). The home visit observations were coded using the HOVRS-A+ (Roggman et al., 2010). The HOVRS-A+ contains seven scales to measure the quality of a home visit: home visitor responsiveness to family (Cronbach's alpha = .75), home visitor-family relationship (Cronbach's alpha = .89), home visitor facilitation of parent-child interaction (Cronbach's alpha = .82), home visitor non-intrusiveness (Cronbach's alpha = .87), parent-child interaction (Cronbach's alpha = .92), parent engagement (Cronbach's alpha = .83), and child engagement (Cronbach's alpha = .91).

In the present study, Home Visitor Practices for home visitors in the intervention group were within the adequate-to-good quality range (M = 3.79, SD = 0.97), while the comparison group scored significantly lower (M = 2.77, SD = 1.04), t(39) = 3.04, p = .004. Family Engagement was good for both the intervention (M = 5.02, SD = 1.39) and comparison (M = 4.48, SD = 1.38) groups. These overall scores are similar to previous studies that used the HOVRS-A+. For example, the Illinois Prevention Initiative Monitoring (Korfmacher et al., 2012) found that the Home Visitor Practices were within the adequate-to-good quality range (M = 3.71, SD = 1.03), as was Family Engagement (M = 4.59, SD = 1.11).

Because the data were nested—that is, families are nested within home visitors, and home visitors are nested within programs—we analyzed the three-level hierarchical linear models (programs, home visitors, and families) for each of the dependent variables in the home visiting observation data, which are the seven scales of the HOVRS-A+. Table 25 displays the sample: a total of 7 home visitors and 42 families participated in home visit observations. The same home visitors participated over time, while different families participated in each observation.

Table 25. Home Visit Observation Sample Size by Group and Time (N = 7 Home Visitors and 42 Families)

	Baseline	Time 2	Time 3	Time 4	Total
Intervention					
Home Visitors	3	3	2	2	3
Families	6	4	3	2	15
Comparison					
Home Visitors	4	4	4	3	4
Families	8	8	6	5	27

After examining descriptive statistics of our sample, we removed variables that were not statistically significant or that were highly correlated with other variables. The following variables were tested in the models predicting the home visit observations:[31]

- Staff age (2 variables)
- Staff race and ethnicity (4 variables)
- Staff educational attainment (2 variables)
- Parent age (2 variables)
- Parent race and ethnicity (4 variables)
- Parent educational attainment (2 variables)
- Staff burnout (MBI - Emotional Exhaustion)
- Staff depression (PHQ-2)
- Staff reflective capacity (RFQ - Certainty)
- Staff relationship with supervisors (SWAI - Rapport)
- Staff self-efficacy (TOS and GAS)
- Duration of family enrollment in program
- Intervention or comparison group
- Time

The intervention group had significantly higher scores on two of the HOVRS-A+ scales: the Home Visitor Responsiveness to Family scale and the Home Visitor Facilitation of Parent-Child Interaction scale. See Table 26 for results of the LMM analyses for the final model for each scale on the HOVRS-A+. Table C-2 in the appendix includes the descriptive statistics for the HOVRS-A+ scales.

The Home Visitor Responsiveness to Family scale score was significantly higher for the home visitors in the intervention group than the comparison group, meaning that the home visitors who received the intervention were more frequently engaged in responsive behaviors during the home visit and elicited input on the content and activities of the home visit from the parent (Roggman et al., 2010).

Figure 12 displays the mean scores on this scale for home visitors in each group at each observation time. In addition, the estimated fixed effect of time on the overall score of Responsiveness to Family scale was positive, indicating that home visitors' scores on this scale tended to increase over time.

[31] Details on the analytic approach and the variables included in the models can be found in Appendix A.

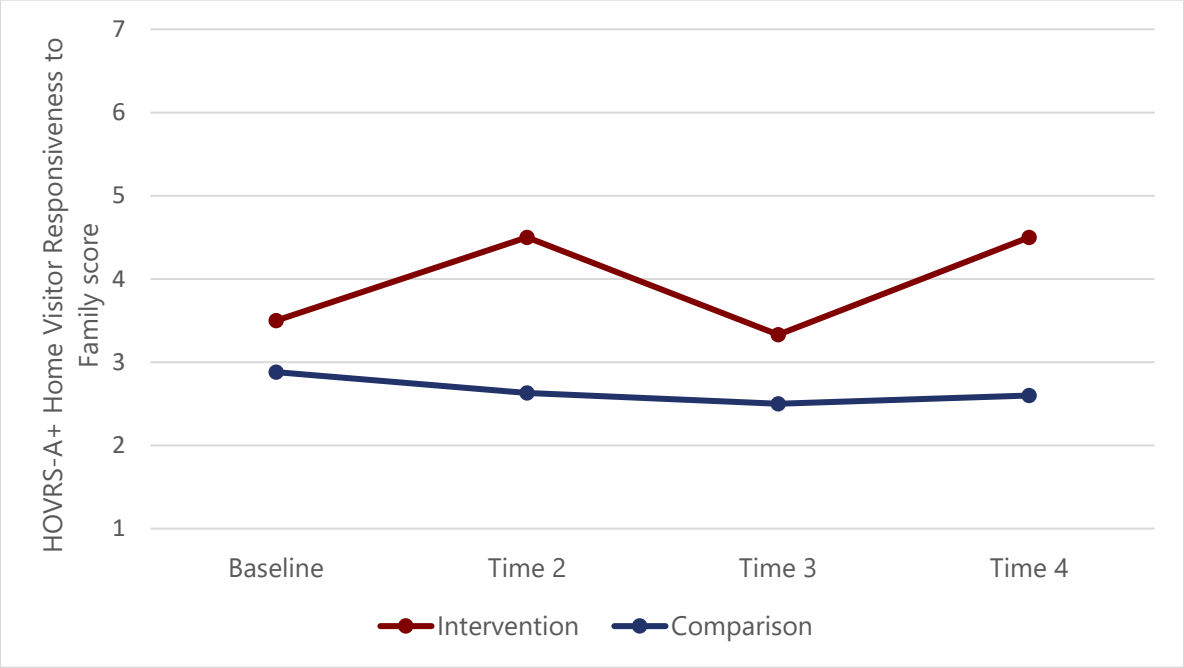
Table 26. Home Visit Observations: LMM Analysis Results (*N* = 7 home visitors, 42 families)

HOVRS-A+ scale	Intervention effect	Effect of any other variables
Home Visitor Responsiveness to Family	β = 1.26, *p* = .044*	Time, β = 1.09, *p* = .083^
Home Visitor–Family Relationship	β = 0.91, *p* = .359	N/A
Home Visitor Facilitation of Parent-Child Interaction	β = 1.11, *p* = .080^	Parent age, β = 0.73, *p* = .029* Parent Hispanic ethnicity, β =-0.97, *p* = .027*
Home Visitor Non-Intrusiveness/Collaboration with Family	β = 0.72, *p* = .222	Parent race (Black), β = 0.88, *p* = .045*
Parent-Child Interaction during Home Visit	β = 0.87, *p* = .188	Parent language (Spanish), β = -1.84, *p* = .004** Parent education, β = -0.99, *p* = .030* Duration of enrollment, β = 0.02, *p* = .094^
Parent Engagement during Home Visit	β = 0.53, *p* = .400	Parent language (Spanish), β = -0.95, *p* = .021*
Child Engagement during Home Visit	β = 0.58, *p* = .329	Parent language (Spanish), β = -1.59, *p* = .006** Time, β = 3.17, *p* = .0565^ Child age, β = 0.06, *p* = .003**

^*p* < .10, **p* < .05, ***p* < .01

Note. Our usual p-level for statistical significance is < .05. However, we also report results of p < .10 to show a trend in the data.

Figure 12. Home Visitor Responsiveness to Family by Group over Time (*N* = 7 home visitors, 42 families)

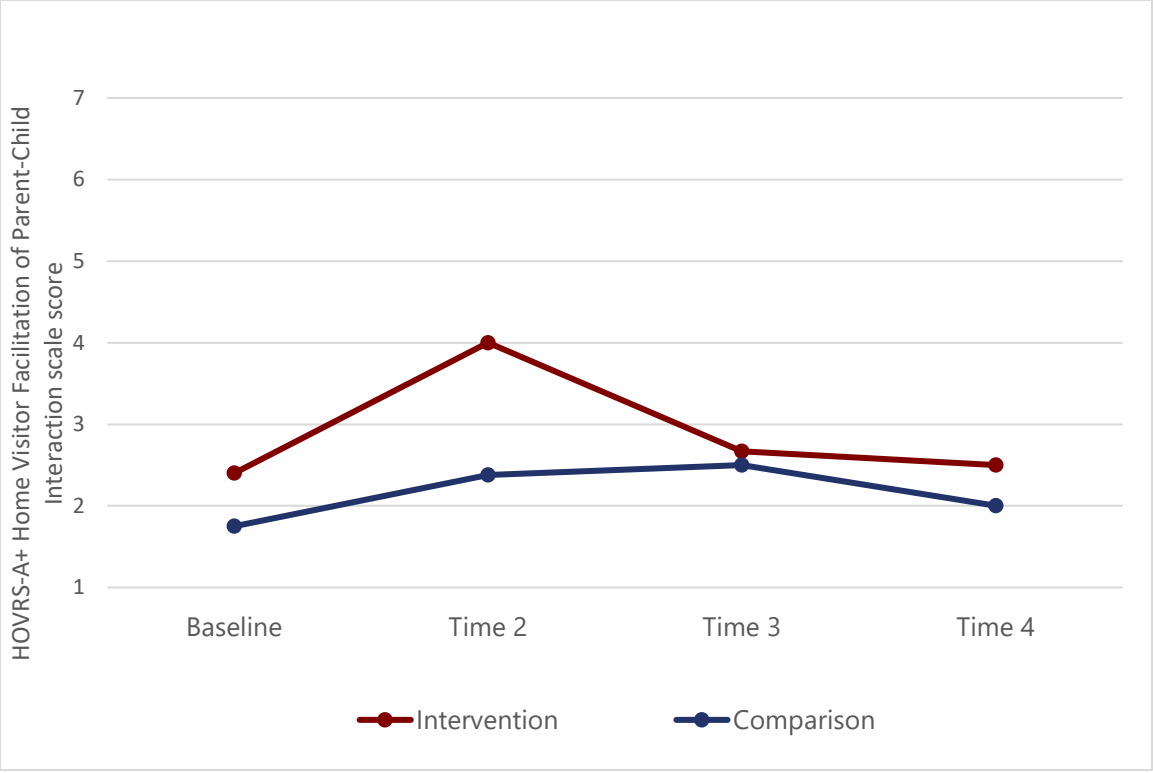


Note. Main effect of the intervention was significant on the Home Visitor Responsiveness to Family scale, β = 1.26, *p* = .044.

Home visitors in the intervention group scored higher, on average, than those in the comparison group on the Home Visitor Facilitation of Parent-Child Interaction scale. This means that home visitors in the intervention group facilitated and promoted positive parent–child interactions and encouraged the parent's leadership in interactions during the visit more often than home visitors in the comparison group.

Figure 13 displays the mean scores on this scale by group at all four observation times. The estimated fixed effect of the categorical variable of parent age on the overall score scale was also positive, which means that home visitors working with parents 30 years old or older had higher overall scores for this scale than home visitors working with younger parents. Finally, the estimated fixed effect of Hispanic parent ethnicity on the overall score of the Home Visitor Facilitation of Parent-Child Interaction scale was negative, suggesting that home visitors working with Hispanic parents have lower overall scores for this scale than home visitors working with White or Black parents.

Figure 13. Home Visitor Facilitation of Parent-Child Interaction (*N* = 7 home visitors, 42 families)



Note. Main effect of the intervention on the Home Visitor Facilitation of Parent-Child interaction scale, β = 1.11, p = .080.
Our usual p-level for statistical significance is < .05. However, we also report results of p < .10 to note a trend in the data.

Although the intervention did not have a significant effect on the other HOVRS-A+ scales, we did find significant effects of other variables. The estimated fixed effect of Black parent race on the overall score of the Home Visitor Non-Intrusiveness/Collaboration with Family scale was positive, meaning that home visitors working with Black parents had higher overall scores for this scale than home visitors working with White or Hispanic parents. Thus, home visitors working with Black parents tend to be less intrusive in a manner that promotes collaboration with the parent as a partner in supporting the child's development.

Spanish parent language had a significant negative effect on three scales of the HOVRS-A+ measuring family engagement: Parent-Child Interaction during the Home Visit, Parent Engagement during the Home Visit, and Child Engagement during the Home Visit. This suggests that Spanish-speaking parents in this sample were less frequently responding to and interacting with their children during visits, less involved in visits, and their children were less engaged and interested in the visits than English-speaking parents. Parent educational attainment also had a negative effect on their score on the Parent-Child Interaction scale during the Home Visit scale, suggesting that parents who completed more than high school were less frequently engaging in warm, positive behaviors during the home visit, compared to parents who completed high school or less. In other words, parents with less education had more positive parent-child interactions. The length of time the family was enrolled in the program also had an effect on the overall score of the Parent-Child Interaction during the Home Visit scale, in that the longer a family was enrolled in the home visiting program, the more frequently they engaged in warm, positive behaviors during the home visit.

For the Child Engagement during the Home Visit scale, the estimated fixed effect of time on the overall score was positive, suggesting that, on average, overall scores for this scale increase over time. (Note that because different families participated in each observation, this should not be interpreted as the child becoming more engaged over time. Rather, the home visitor's practice may be changing in a manner that facilitates more child engagement over time.) In addition, the estimated fixed effect of child age on the overall score of the Child Engagement during the Home Visit scale is small but positive, suggesting that older children are more frequently displaying behaviors that indicate engagement and interest in the home visit, compared to younger children.

IECMHC Scale

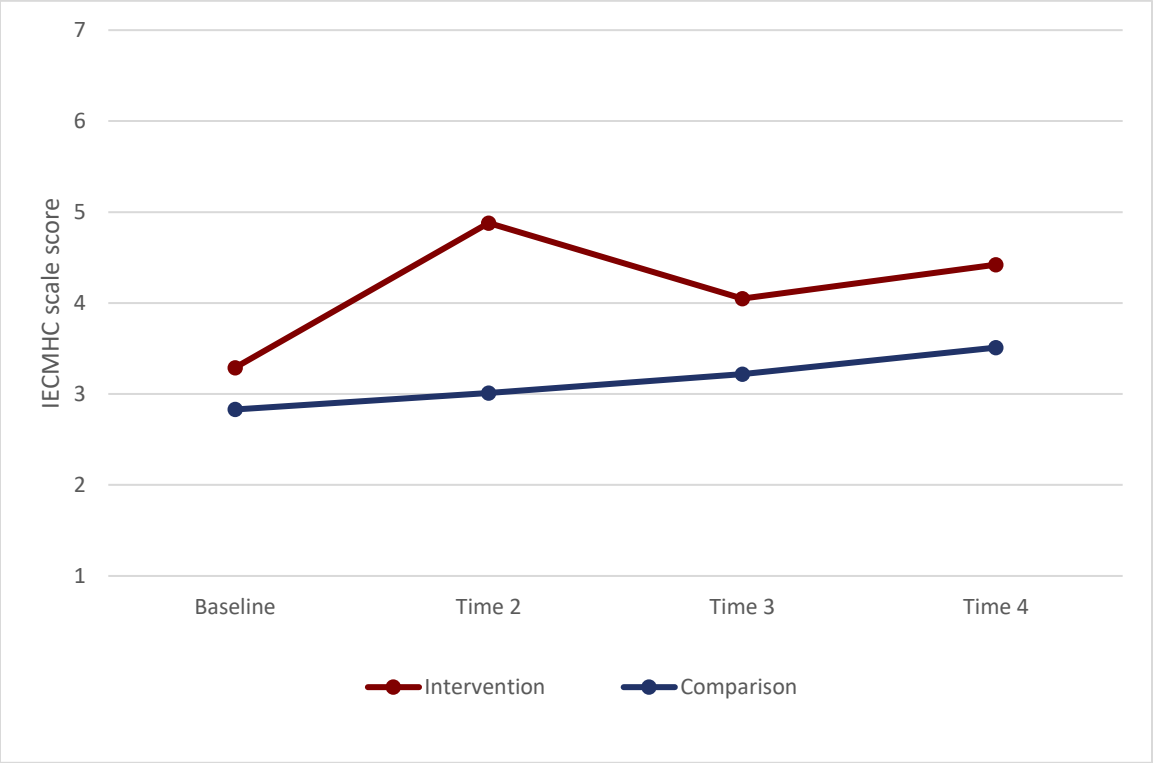
The HOVRS-A+ was not designed specifically to measure effects of IECMHC. Therefore, we created a scale comprised of the HOVRS-A+ items that we expected would be impacted by the intervention, according to the Illinois Model's theory of change.[32] Cronbach's alpha for this scale

[32] The items contained in this IECMHC scale are the following items from the HOVRS-A+: Responsiveness to Family 3, Responsiveness to Family 4, OVERALL scale score for Relationship with Family, Facilitation of

was .92, indicating high reliability. We conducted a nested two-way ANOVA which tests for differences between the intervention and comparison groups over time, with staff nested in group to account for stronger correlations within the visits for each home visitor (i.e., each home visitor's visits will correlate to their own visit scores more than they will correlate to each other's scores). The interaction term was significant, $F(24, 40) = 2.31$, $p = .044$, $\eta_p^2 = .78$. The home visitors in the intervention group had scores on the IECMHC items in the HOVRS-A+ that tended to increase at a greater rate than the scores of the home visitors in the comparison group.

Figure 14 displays the IECMHC scale means by group over time.

Figure 14. IECMHC items in Home Visit Observations (*N* = 7 home visitors, 42 families)



Note. Two-way interaction effect between time and group on IECMHC items in the HOVRS-A+ was significant, F (24, 40) = 2.31, p = .044.

Child and Family Outcomes

The Illinois Model of IECMHC is designed to support the mental health of children and families by strengthening the knowledge, skills, and reflective capacity of the early childhood workforce. Thus, this intervention focused on staff, rather than children and families. At the same time,

PCI 1, Facilitation of PCI 2, Facilitation of PCI 4, Facilitation of PCI 5, Nonintrusiveness/Collaboration 4, Nonintrusiveness/Collaboration 5, Parent-Child Interaction 4, Parent-Child Interaction 5, Parent-Child Interaction 6, Parent-Child Interaction 7. Cronbach's alpha for this scale was .92.

according to the theory of change, improvements in staff capacity should positively affect children and families.

Child Outcomes: Teachers' Assessments of Children's Behaviors

Intervention and comparison group teachers completed assessments on a small group of children (up to 8 per focus sample classroom) at baseline, Time 2, and Time 3. The sample of 136 children—83 in the intervention group and 53 in the comparison group—were assessed by 21 teachers, representing 12 programs. See Table 27 for sample characteristics of the children who were assessed. Race and ethnicity data were collected at Time 2, so we can only report race/ethnicity data for the 55 children who were assessed at all three time points. However, we have gender and age data for all children who were assessed at baseline.

Table 27. Child Assessments Sample Characteristics (*N* = 136)

	Intervention (*n* = 83)	Comparison (*n* = 53)
Gender		
Female	50.6%	43.4%
Male	49.4%	56.6%
Age	Mean = 51.7 months (*SD* = 8.05)	Mean = 54.2 months (*SD* = 9.48)
Race and Ethnicity		
Black	9.3%	33.3%
Hispanic	44.2%	58.3%
White	46.5%	8.3%

Note. Child gender, age, race, and ethnicity were reported by teachers

Three standardized measures were used to assess child behavior and development: the Strengths and Difficulties Questionnaire (SDQ) Impact Supplement (Goodman, 1997; brief version, Perry, 2013; Stephan et al., 2011), Preschool Expulsion Risk Measure (PERM; Gilliam & Reyes, 2018), and the Devereux Early Childhood Assessment (DECA-P2; LeBuffe & Naglieri, 2012) version for children ages 36–60 months.

- The SDQ Impact Supplement asks the teacher whether the child has difficulties in one or more of the following areas: emotions, concentration, behavior, or being able to get along with other people. For children were rated as having any difficulties, additional questions are asked about distress, social impairment, and learning, resulting in an Impact Score. Cronbach's alphas for this scale were .65, .67, and .75 at baseline, Time 2, and Time 3, respectively.
- The PERM measures the likelihood that a program would consider expelling a child, providing a total score and four subscales:
    - Classroom Disruption (Cronbach's alphas were .95, .96, .93 at baseline, Time 2, and Time 3, respectively)
    - Accountability (Cronbach's alphas were .90, .81, .81 at baseline, Time 2, and Time 3, respectively)

    - Hopelessness (Cronbach's alphas were .79, .86, .84 at baseline, Time 2, and Time 3, respectively)
    - Teacher Stress (Cronbach's alphas were .94, .88, .93 at baseline, Time 2, and Time 3, respectively).
- The DECA-P2 includes three scales assessing protective factors:
    - Initiative (Cronbach's alphas were .89, .93, .93 at baseline, Time 2, and Time 3, respectively)
    - Self-Regulation (Cronbach's alphas were .95, .94, .93 at baseline, Time 2, and Time 3, respectively)
    - Attachment/Relationships (Cronbach's alphas were .81, .79, .82 at baseline, Time 2, and Time 3, respectively), summed to provide a Total Protective Factors scale score
- The DECA-P2 also includes a Behavioral Concerns scale (Cronbach's alphas were .87, .85, .83 at baseline, Time 2, and Time 3, respectively)

After examining descriptive statistics of our sample, we removed variables that were not statistically significant or that were highly correlated with other variables. The following variables were tested in the models for the child assessments:[33]

- Teacher age (2 variables)
- Teacher race and ethnicity (4 variables)
- Teacher educational attainment (2 variables)
- Teacher years of experience (1 variable)
- Child race and ethnicity (4 variables)
- Child gender (1 variable)
- Child age (1 variable)
- Teacher burnout (MBI - Emotional Exhaustion)
- Teacher depression (PHQ-2)
- Teacher reflective capacity (RFQ - Certainty)
- Teacher relationship with supervisors (SWAI - Rapport)
- Teacher self-efficacy (TOS)
- Classroom mental health climate (CHILD Overall Score)
- Intervention or comparison Group
- Time

The descriptive statistics for the child assessment scales can be found in Table C-3 in the appendix. See Table 28 for results of child assessment models.

[33] Details on the analytic approach and the variables included in the models can be found in Appendix A.

Table 28. Child Assessments: LMM Analysis Results (*N* = 136 children in 21 classrooms[a])

Scale	Intervention effect (Time*Group)	Effect of any other variables
Strengths & Difficulties Questionnaire – Impact Score (*n* = 94)	β = -0.42, *p* = .087^	Child gender (female), β = -1.17, *p* = .002** Teacher age, β = 0.78, *p* = .061^
Preschool Expulsion Risk Measure – Total Score (*n* = 117)	β = -0.047, *p* = .81	Child gender (female), β = -0.55, *p* = .001**
Preschool Expulsion Risk Measure – Classroom Disruption (*n* = 44)	β = 0.064, *p* = .92	Child gender (female), β = -0.65, *p* = .034*
Preschool Expulsion Risk Measure – Fear of Accountability	β = -0.18, *p* = 0.285	Child gender (female), β = -0.73, *p* < .001*** Teacher years of experience, β = 0.032, *p* = .011*
Preschool Expulsion Risk Measure – Hopelessness (*n* =136)	β = -0.15, *p* = .155	Child gender (female), β =-0.40, *p* = .001** Teacher age, β = 0.29, *p* = .024*
Preschool Expulsion Risk Measure – Teacher stress (*n* = 110)	β = 0.021, *p* = .914	Teacher reflective capacity, β = -0.29, *p* = .047*
DECA – Total Protective Factors (*n* = 119)	β = 0.59, *p* = .957	Time, β = 50.03, *p* = .030* Child gender (female), β = 12.82, *p* = .001** Teacher age, β = -6.701, *p* = .083^ Teacher reflective capacity, β = 6.26, *p* = .063^ Teacher self-efficacy*Time, β = -0.79, *p* = .056^
DECA – Initiative (*n* = 96)	β = -1.64, *p* = .511	Time, β = 27.46, *p* = .001** Teacher self-efficacy*Time, β = -0.33, *p* = .017*
DECA – Self-Regulation (*n* = 122)	β = 0.19, *p* = .929	Child gender (female), β = 6.46, *p* < .001*** Teacher depression, β = -1.73, *p* = .053^
DECA – Attachment/ Relationships (*n* = 122)	β = 1.21, *p* = .776	Child gender (female), β = 3.63, *p* = .009**
DECA – Behavioral Concerns (*n* = 98)	β = 1.07, *p* = .696	Child gender (female), β = -8.41, *p* < .001*** Classroom mental health climate, β = -2.08, *p* = .081^ Teacher depression, β = 1.91, *p* = .017* Teacher reflective capacity, β = -2.86, *p* = .032*

^*p* < .10, **p* < .05, ***p* < .01, ****p* < .001.

Note. Our usual *p*-level for statistical significance is < .05. However, we also report results of *p* < .10 to show a trend in the data.

[a] We could only follow a sample of 55 children over time, but the LMM analysis includes the full baseline sample of 136.

Strengths and Difficulties Questionnaire Impact Supplement. In the brief version of the SDQ Impact Supplement (Perry, 2013; Stephan et al., 2011) used in this study, the first item measures the presence of any child behavioral or emotional difficulties. For the children who are rated as having any difficulties, additional questions are asked about whether these difficulties upset the

child, interfere with the child's peer relationships, or interfere with the child's learning, resulting in an Impact Score. The effect of the intervention was trending toward significant over time on the SDQ Impact Score. Children in classrooms receiving the intervention had a lower average Impact Score over time than children in comparison classrooms, after controlling for the effects of the other variables in the model. The Impact Score items are only administered when the child is rated as having any behavioral difficulties. Thus, according to their teachers, the behavior problems of children in the intervention group who had any behavioral problems tended to have less severe behavioral problems over time than children in the comparison group.

There was also a negative effect of student gender (female) on the SDQ Impact Score, after controlling for the effects of the other variables in the model. Girls had a lower average impact score than boys. Thus, teachers perceived girls' behavioral problems to be less severe than boys' behavioral problems. In addition, there was a trend of a positive effect of staff age on the SDQ Impact Score, after controlling for the effects of the other variables in the model. Older staff tended to rate children's behavioral severity higher than younger staff.

Preschool Expulsion Risk Measure (PERM). The intervention did not have a significant effect on PERM total score or any of the four PERM subscales. There was, however, a negative effect of student gender (female) on the PERM total score, after controlling for the effects of the other variables in the model. Girls had a lower average PERM total score than boys. Thus, according to their teachers, girls were at lower risk of expulsion than boys.

For the Classroom Disruption subscale, there was a negative effect of child gender (female), after controlling for the effects of the other variables in the model. Girls had a lower average score on the Classroom Disruption subscale than boys, meaning their teachers perceived that girls were less likely to create disruptions in the classroom than boys.

The Fear of Accountability subscale measures the degree to which children's behaviors may pose a risk of injury for which the teacher might be accountable. There was a positive effect of teacher years of experience on the Fear of Accountability subscale, after controlling for the effects of the other variables in the model. Teachers with more years of experience working with children had a higher average score on this subscale than teachers with less experience, indicating that more experienced teachers were more likely to view children's behaviors as posing a risk of injury for which they might be accountable. There was also a negative effect of child gender (female) on the Fear of Accountability subscale, after controlling for the effects of the other variables in the model. Girls had lower average scores than boys on this subscale, suggesting that girls were less likely than boys to engage in behaviors that might pose a risk of injury for which their teachers might be accountable, according to their teachers.

The Hopelessness subscale on the PERM measures the degree to which the teacher feels hopeless that anything can be done to improve behaviors in the classroom. There was a positive

effect of teacher age on the Hopelessness subscale, after controlling for the effects of the other variables in the model. Children in classrooms led by older staff had a higher average score on the Hopelessness subscale than children in classrooms led by younger staff. Thus, older staff felt more hopeless that anything can be done to improve behaviors in the classroom than younger staff. There was also a negative effect of child gender (female) on the Hopelessness subscale, after controlling for the effects of the other variables in the model. Girls had lower average scores than boys on this subscale, indicating that teachers felt more hopeless about improving boys' behavior than they did about improving girls' behavior.

Although the intervention did not have a significant effect on the PERM Teacher Stress subscale, teacher reflective capacity had a significant negative effect on teacher stress, after controlling for the effects of the other variables in the model. Children in classrooms led by teachers with higher reflective capacity had a lower average score on the PERM Teacher Stress subscale compared to children in classrooms led by teachers with lower reflective capacity. This means that teachers with higher reflective capacity reported less teacher stress associated with children's behaviors than teachers with lower reflective capacity.

Devereux Early Childhood Assessment (DECA). The effect of the intervention was not significant on any scales in the DECA, but several other variables—child gender and age; teachers' age, depression, and reflective capacity; and time—did affect scores on the DECA scales. The Total Protective Factors scale on the DECA—comprised of the Initiative, Self-Regulation, and Attachment subscales—gives an overall indication of the child's social and emotional strengths related to resilience (LeBuffe & Naglieri, 2012). There was a positive overall effect of time on the Total Protective Factors scale, after controlling for the effects of the other variables in the model. Thus, children in both comparison and intervention classrooms showed an increase over time on social and emotional strengths related to resilience.

There was also a positive effect of child gender (female) on the Total Protective Factors scale, after controlling for the effects of the other variables in the model. Girls had a higher average standard score, demonstrating greater social and emotional strengths related to resilience, compared to boys. There was a negative effect of teacher age on the Total Protective Factors scale, after controlling for the effects of the other variables in the model. Children in classrooms led by older staff had a lower average standard score than children in classrooms led by younger staff. Thus, younger teachers viewed children in their classrooms as having more social and emotional strengths related to resilience than older teachers.

Teacher reflective capacity (RFQ – Certainty) scores had a positive effect on the Total Protective Factors scale, after controlling for the effects of the other variables in the model. Children in classrooms led by teachers with higher reflective capacity had higher average standard scores than children in classrooms led by teachers with lower reflective capacity. Thus, teachers with

greater reflective capacity perceived children in their classroom to have more social and emotional strengths than teachers with lower reflective capacity.

Teacher self-efficacy (TOS score) had an inverse relationship with Total Protective Factors scale scores over time, after controlling for the effects of the other variables. Children in classrooms led by teachers with higher self-efficacy scores had lower scores over time on the Total Protective Factors scale than children in classrooms led by teachers with lower self-efficacy scores. Thus, teachers with higher self-efficacy perceived children in their classrooms had weakened social and emotional strengths related to resilience over time.

Time had a positive overall effect on the DECA Initiative scale, after controlling for the effects of the other variables in the model. On average, children's scores on the Initiative scale increased over time. This means that over time, teachers perceived children in both groups (intervention and comparison) to be better able to use independent thought and action to meet their needs.

Similar to the Total Protective Factors scale, there was a negative effect of time on the DECA Initiative scale for teachers with higher self-efficacy scores, after controlling for the effects of the other variables in the model. Children in classrooms led by teachers with higher self-efficacy had lower average Initiative scores over time than children in classrooms led by teachers with lower self-efficacy scores. Thus, teachers with higher self-efficacy tended to perceive children in their classrooms to be less able to use independent thought and action to meet their needs over time compared to children in classrooms of teachers with lower self-efficacy.

The DECA Self-Regulation scale measures the child's ability to express and manage emotions without exhibiting challenging behaviors. Child gender (female) had a positive effect on the Self-Regulation scale score, after controlling for the effects of the other variables in the model. Girls had higher Self-Regulation scale scores than boys; therefore, teachers generally reported that girls were more able to express emotions and manage behavior in healthy ways than boys. In addition, teacher depression had a negative effect on child self-regulation. Children in classrooms led by teachers with higher depression had lower scores on the Self-Regulation scale than children in classrooms led by teachers with lower depression. Thus, teacher depression was associated with a more negative view of children's ability to effectively manage their emotions and behavior.

The Attachment/Relationships subscale on the DECA measures the child's ability to promote and maintain mutual, positive connections with other children and significant adults. Child gender (female) had a positive effect on the Attachment/Relationships scale score, after controlling for the effects of the other variables in the model. Girls were rated by their teachers as better able than boys to promote and maintain mutual, positive connections with other children and significant adults.

The DECA Behavioral Concerns scale assesses challenging behaviors in children, including aggression, withdrawal, inattention, and exhibiting extreme emotions. Once again, child gender had a significant effect on this scale, in that girls were rated by their teachers as having fewer behavioral concerns than boys. The overall score on the classroom observation measure, the CHILD, also had a negative effect on the DECA Behavioral Concerns scale. Thus, children in classrooms that promote a positive mental health climate had fewer teacher-reported problems with aggression, withdrawal, attention, and the control of extreme emotions.

Teacher depression had an effect on children's behavioral concerns, after controlling for the effects of the other variables in the model. Children in classrooms led by teachers with higher depression scores had higher Behavioral Concerns scale scores, on average, than children in classrooms led by teachers with lower depression scores. Thus, teacher depression scores predicted teacher-reported child internalizing and externalizing behavioral concerns.

There was a negative effect of staff reflective capacity scores on the DECA Behavioral Concerns scale, after controlling for the effects of the other variables in the model. Children in classrooms led by teachers with higher reflective capacity scores had lower scores on this scale than children in classrooms led by teachers with lower reflective capacity scores. Thus, teachers with higher reflective capacity perceived the children in their classrooms to have fewer problems with aggression, withdrawal, attention, and the control of extreme emotions than teachers with lower reflective capacity.

Family Outcomes: Healthy Parenting and Parent Well-being

Parents who participated in the video-recorded home visit observations completed a survey. The parent survey included the Patient Health Questionnaire-2 (PHQ-2; Kroenke, Spitzer, & Williams, 2003) to measure depression and five subscales from the Healthy Families Parenting Inventory (HFPI; LeCroy & Milligan Associates, 2004): Problem Solving, Parent/Child Interaction, Home Environment, Role Satisfaction, and Parenting Efficacy. Parents with infants under 12 months of age were also administered the Infant Crying & Parent Well-being screening tool (Katch & Burkhardt, 2021) to assess their perception of infant crying and self-efficacy related to soothing.

Because home visitors recorded visits with different families at each data collection point, different parents completed the parent survey each time.

Table 29 presents the parent demographic characteristics. All parent participants identified themselves as female. The target child for the home visiting program was an average of 19 months of age (M = 19.67, SD = 13.03), but ranged from 4 months to 50 months of age. The number of children in the household ranged from 1 to 5, with a mean of 2.59 (SD = 1.19) children. Chi-square tests revealed no significant group differences in language or education in the parent survey sample. There were significant differences, however, between parent survey participants in the intervention and comparison groups in the distribution of parent age, χ^2 (3, N

= 51) = 11.20, p = .011, and parent race and ethnicity, χ^2 (2, N = 50) = 15.21, p < .001. Parents in the intervention programs were younger and more likely to be non-Hispanic White, compared to parents in comparison programs.

After reviewing descriptive statistics (see Table C-4 in the appendix), we tested for covariates by conducting correlational analyses (for continuous variables) and ANOVAs (for categorical variables) with demographic characteristics and the parent survey scales (HFPI scales and PHQ score). There were no significant differences on the parent survey scale scores by parent education, parent age, child age, or number of children in the family. There was a significant difference by race and ethnicity for HFPI Parenting Efficacy, F(2, 47) = 5.10, p = .010). Black mothers reported higher parenting efficacy (M = 28.17, SD = 2.20) than White (M = 24.20, SD = 4.61) and Hispanic mothers (M = 25.45, SD = 3.70). Therefore, we included race and ethnicity as a covariate[34] in the analysis of the Parenting Efficacy scale.

Table 29. Parent Survey Sample Characteristics (N = 51)

Characteristic	Intervention n = 24	Comparison n = 27
Preferred language	%	%
English	75.0	53.8
Spanish	25.0	46.2
Age		
18–29	70.8	25.9
30–39	29.2	63.0
40–49	0	7.4
50+	0	3.7
Race/Ethnicity		
Black	30.4	40.7
Hispanic	26.1	59.3
White	43.5	0
Education		
Less than high school	4.2	8.7
Some high school	16.7	8.7
High school/GED	45.8	56.5
Some college	29.2	17.4
Bachelor's degree	4.2	8.7

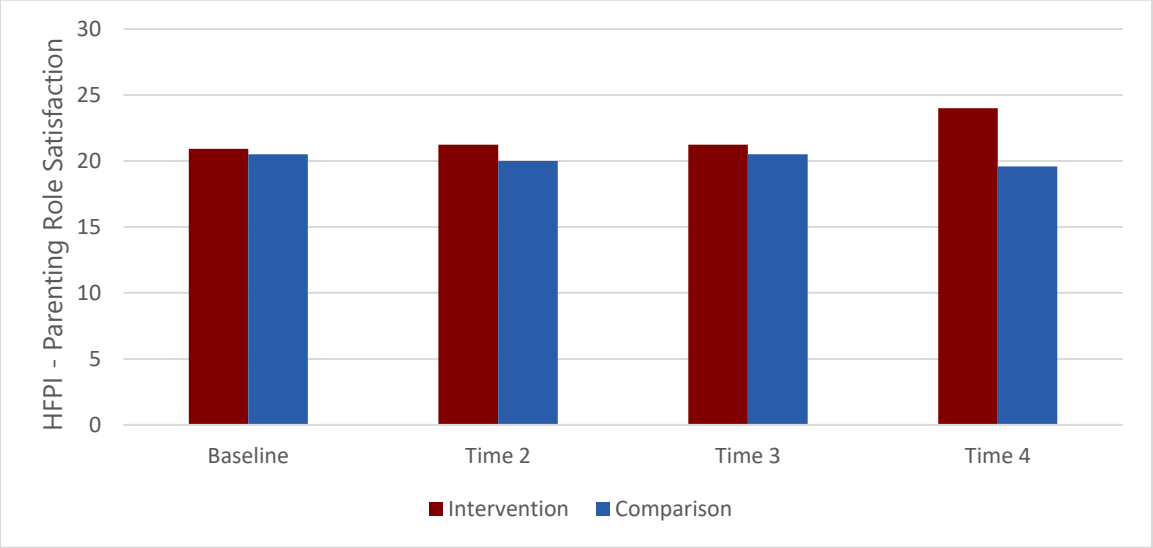
We used a multilevel analysis to account for the fact that some families were participating in the same programs. We conducted hierarchical two-way ANOVAs, with program nested within group over time, to examine whether there were differences between the scores of the parents whose home visitors were receiving the intervention and the parents whose home visitors were

[34] Race/ethnicity was included as a dummy coded variable, with Black = 1, White/Hispanic = 0.

not receiving the intervention. The Problem Solving scale had an interaction effect of group by time that was trending toward significance, $F = 1.82$, $p = .088$, $\eta_p^2 = .41$. However, the mean scores by group over time reveal that the parents in the intervention group had mean scores that were similar at each time point, while the mean scores for the parents in the comparison group fluctuated.

Because the parent sample was small and some home visiting programs had just one parent participant in the evaluation at a time, we also analyzed the parent survey data without the nested analysis. We tested for main effects of the intervention because the data were not repeated at the parent level. At baseline, no significant differences were found between parents in the intervention and comparison groups on any of the measures in the parent survey. We analyzed the parent surveys with a univariate ANOVA, comparing the two groups at all four time points. We found a significant main effect of the intervention on the Role Satisfaction subscale, $F(1) = 4.73$, $p = .035$, $\eta_p^2 = .10$. Thus, parents whose home visitors had been receiving the intervention scored significantly higher on Role Satisfaction and were more comfortable and content with being a parent than the comparison group parents. See Figure 15 for parent survey scores on the Role Satisfaction subscale at each time.

Figure 15. Parent Role Satisfaction by Group (*N* = 51)



Note. The intervention had a significant effect on the HFPI Role Satisfaction subscale, $F(1) = 4.73$, $p = .035$.

Follow-up Survey of Parents

At Time 4, we followed up with parents participating in the home visiting programs who completed a survey at baseline, Time 2, or Time 3. At Time 4, we administered the parent survey to them a second time so we could compare any changes over time within the families between groups. We were able to contact 21 of 35 parents who completed a prior survey and administer

a follow-up survey. Sample characteristics for the parent participants who completed the survey twice are presented in Table 30.

Chi-square tests revealed no significant group differences other than the distribution of parent age, χ^2 (2, N = 21) = 10.94, p = .004, and parent race and ethnicity, χ^2 (2, N = 21) = 6.91, p = .032, between the intervention and comparison groups for parents in the follow-up sample. Parents in the follow-up sample in the intervention group tended to be younger than those in the comparison group, and there was a greater proportion of Hispanic parents in the comparison group follow-up sample, compared to the intervention group. We analyzed the follow-up sample of parent surveys by calculating change scores (subtracting their first survey scores from their follow-up survey scores) and then conducted a t-test to compare the intervention group and the comparison group for each scale's change score. This test analyzed whether any changes over time that occurred were different between parents whose home visitors were in the intervention group versus the comparison group. We did not find any significant group differences (see Table C-8).

Table 30. Follow-up Parent Survey Sample Characteristics (*N* = 21)

	Intervention (*n* = 9)	Comparison (*n* = 12)
Preferred language	%	%
English	66.7	25.0
Spanish	33.3	75.0
Age		
18–29	88.9	16.7
30–39	11.1	83.3
40+	0	0
Race/Ethnicity		
Black	22.2	25.0
Hispanic	33.3	75.0
White	44.4	0
Education		
Less than high school	0	16.7
Some high school	11.1	16.7
High school/GED	55.6	50.0
Some college	33.3	16.7

Home Visit Observations and Parent Outcomes

To examine the relationship between the features observed in the home visits and parent outcomes, we analyzed the HOVRS-A+ scores and the scores from the parent survey measures. Results of the correlation analysis are displayed in Table 31.

Focusing on the HOVRS-A+ scales on which the intervention had an effect, Home Visitor Responsiveness to Family was correlated with Role Satisfaction and Home Environment, Home Visitor Facilitation of Parent-Child Interaction was correlated with Parent-Child Interaction, and the IECMHC scale we developed was also associated with parent's report of their interactions with their child. In addition, Home Visitor Facilitation of Parent-Child Interaction was also negatively correlated with parental depression. Note that these associations do not provide any information on causality, only that a relationship exists.

Table 31. Correlations between Home Visit Ratings and Parent Outcomes (*N* = 42)

	HFPI-Problem Solving	HFPI-Role Satisfaction	HFPI-Parenting Efficacy	HFPI-Parent/Child Interaction	HFPI-Home Environment	Parent depression (PHQ Score)
Home Visitor Responsiveness to Family	0.272^	.343*	0.021	0.196	.344*	-0.110
Home Visitor Relationship with Family	0.266^	0.294^	-0.068	0.179	0.169	-0.041
Home Visitor Facilitation of Parent-Child Interaction	0.265^	0.264^	0.196	.445**	.310*	-0.276^
Home Visitor Nonintrusiveness/ Collaboration	0.010	0.268^	0.130	.391*	0.215	-0.111
Parent-Child Interaction during Home Visit	-0.008	-0.007	.313*	.443**	0.194	-0.079
Parent Engagement during Home Visit	0.103	0.113	0.149	.369*	0.197	-0.123
Child Engagement during Home Visit	0.034	-0.141	0.043	0.281^	0.037	-0.057
Home Visitor Practices domain	0.235	.338*	0.075	.354*	0.296^	-0.152
Family Engagement domain	0.044	-0.029	0.182	.404**	0.149	-0.092
IECMHC items	0.159	0.292^	0.174	.405**	0.294^	-0.128

^p < .10, *p < .05, **p < .01, ***p < .001. Our usual p-level for statistical significance is < .05. However, we also report results of p < .10 to note a trend in the data.

Note. Home Visitor Practices domain includes Home Visitor Responsiveness to Family, Relationship with Family, Facilitation of Parent-Child Interaction, and Nonintrusiveness/Collaboration. Family Engagement domains includes Parent-Child Interaction, Parent Engagement, and Child Engagement.

Chapter Summary

We conducted surveys of staff and supervisors at all programs. In the center-based programs, we observed classrooms and teachers completed child assessments. In the home visiting programs, we rated video recordings of home visits, and participating parents completed surveys. We also conducted interviews with staff, supervisors, directors, and consultants. The

intervention showed significant effects in several outcomes and across data sources. Table C-5 displays a summary of the outcomes.

The intervention showed a significant positive effect on staff reflective capacity on two standardized measures (RFQ in surveys, PRPAS in interviews). Staff at intervention programs in the focused sample were also more likely to report lower levels of burnout by Time 3, compared to staff at comparison programs. Interview data confirmed these findings, along with additional themes of 1) active listening and deeper exploration of issues, 2) the ability to think critically about one's reactions and biases, 3) the ability to consider others' perspectives, and 4) the ability to establish or improve boundaries and be mindful of self-care.

Unexpectedly, the LMM analysis also showed that teacher role was a factor in two measures of the supervisory relationship. Although staff views of their supervisors were fairly positive overall, lead teachers in the intervention group perceived their supervision less favorably than lead teachers in the comparison group. We discuss potential reasons for these findings in the next chapter.

In the center-based programs, there were several significant findings. Classroom observations revealed significant group differences over time. Teachers in the intervention group had a higher average score than comparison classrooms that was statistically significant on the Directions and Rules domain and Equity auxiliary dimension, and that was trending toward significant on the Individualized and Developmentally Appropriate Pedagogy domain. Furthermore, children in classrooms receiving the intervention tended to have a lower average SDQ Impact Score over time than children in comparison classrooms. Although this finding was a trend that did not reach statistical significance, according to their teachers, children with behavioral problems in the intervention group had less severe behavioral problems over time than children in the comparison group.

At the same time, teachers' ratings of behavior showed significant differences by child gender on all child assessments: Teachers consistently viewed boys as having more severe behavioral problems than girls. Notably, teachers rated girls significantly differently than boys on the following domains (measures noted in parentheses):

- Girls' behavioral problems were less severe than boys' behavioral problems (SDQ Impact Score)
- Girls were at lower risk of expulsion than boys (PERM Total Score)
- Girls were less likely to create disruptions in the classroom than boys (PERM Classroom Disruption scale)
- Girls were less likely than boys to engage in behaviors that may pose a risk of injury for which their teachers might be accountable, according to their teacher (PERM Fear of Accountability scale)

- Teachers felt more hopeless that anything can be done to improve boys' behavior compared to girls' behavior (PERM Hopelessness scale)
- Girls were rated as demonstrating greater social and emotional strengths related to resilience compared to boys (DECA Total Protective Factors score)
- Girls were more able to express emotions and manage behavior in healthy ways than boys (DECA Self-Regulation scale)
- Girls were rated by their teachers as better able than boys to promote and maintain mutual, positive connections with other children and significant adults (DECA Attachment/Relationships scale)
- Girls were rated by their teachers as having fewer behavioral concerns than boys (DECA Behavioral Concerns scale)

Some of the child assessments were also affected by teacher age. Specifically, there was a positive effect of staff age on the SDQ Impact Score, after controlling for the effects of other variables in the model. Older staff rated the severity of children's behavioral problems higher than younger staff. There was a negative effect of teacher age on the DECA Total Protective Factors scale, after controlling for the effects of other variables in the model. Children in classrooms led by older staff had a lower average standard score than children in classrooms led by younger staff. Thus, younger teachers viewed children in their classrooms as having more social and emotional strengths than older teachers. In addition, there was a positive effect of teacher age on the Hopelessness subscale, after controlling for the effects of other variables in the model. Children in classrooms led by older staff had a higher average score on the Hopelessness subscale than children in classrooms led by younger staff. Thus, older staff felt more hopeless that children's behaviors in the classroom could improve than younger staff.

We found significant intervention effects in the home visiting programs as well. In the home visit observations, home visitor practice was significantly greater for the intervention group than the comparison group on the Home Visitor Responsiveness to Family scale. The same pattern was found on the Home Visitor Facilitation of Parent-Child Interaction scale, although the effect of the intervention did not quite reach statistical significance. In addition, home visitors in the intervention group scored higher than the comparison group on a scale comprised of HOVRS-A+ items related to IECMHC. In the parent survey, we found a significant main effect of the intervention on the Role Satisfaction subscale (HFPI). Parents whose home visitors were receiving the intervention tended to report higher satisfaction in their role as parents than parents whose home visitors were not receiving the intervention.

Discussion

> I think our time with [consultation] has made a huge impact. It's just having that special time, and knowing you have that time helps carry things that might get heavy through the week. . . and knowing that you're going to have that available to you to decompress, to unpack it, to have somebody that will support you in coming up with some ideas. —*Program Director*

The pilot study of the Illinois Model of IECMHC was conducted with 24 publicly funded early childhood programs—school-based pre-K, Head Start, community-based childcare, and home visiting programs—in urban, suburban, and rural communities in two different regions of the state. We matched programs by location and type and then randomly assigned 16 programs to an intervention group, which received 21 months of services from professional mental health consultants using the Illinois Model. The other eight programs, matched by type, served as a "business as usual" comparison group. Although some of the comparison programs received support from mental health consultants as part of their existing programs, none received services comparable to the Illinois Model.

The evaluation of the Illinois Model of IECMHC required a complex but rigorous design that was flexible and responsive to the community and program characteristics of the sample. Like most approaches to IECMHC, the form and content of consultation in the Illinois Model depends on several factors. These include the needs and goals of the program staff; the relationship between the consultant and staff; program goals, funding, and structure; the stability of staff and organizational leaders; and their readiness to work with a mental health consultant. This made studying implementation—and the fidelity of implementation—complicated.

Although the Illinois Model is based on prior research and tools developed by the field, it is unique in its approach, particularly in the extent to which it emphasizes the development of reflective capacity in staff and supervisors in order to work effectively with children and families. The Illinois Model is the first model of IECMHC designed to work with a range of early childhood systems, including home visiting, community-based childcare, and pre-K programs. In turn, this evaluation was the first to study IECMHC in multiple systems. Thus, it required a unique design and a mix of quantitative and qualitative methods.

In addition to assessing the Illinois Model, this evaluation fills some important gaps in the literature. It provides more in-depth information about the process and challenges of implementing mental health consultation in early childhood systems, and a deeper understanding of the mechanism of change through which IECMHC impacts outcomes. In this chapter, we summarize our key findings by research question and then discuss the strengths

and limitations of the study. We conclude by discussing implications of the findings for policy, practice, and further research.

Summary of Findings

Research Question 1: Was the Illinois Model of IECMHC implemented as intended? What factors affected its implementation?

The Illinois Model was successfully implemented in a variety of early childhood center-based and home visiting programs as measured by structural and process indicators of fidelity. At the same time, there were a number of program factors that affected implementation.

Structural indicators. One of the primary structural indicators for assessing implementation was dosage, or the number of hours of consultation received. Analysis of data from consultant logs indicated that all but two of the programs received at least 80% of their expected consultant goal hours. Another structural indicator is adherence, or the extent to which consultants' activities were consistent with the model. Analysis of the consultant logs showed that although consultant activities varied, all intervention programs received the expected type of consultant support. The most frequent activities were participating in reflective supervision sessions with individual staff and their supervisors; reflective consultation with directors and supervisors; and reflective consultation with staff (without the supervisor present). There were differences between early childhood center-based programs and home visiting programs in types of activities; and there was considerable variability in activities among the programs in each group, reflecting the flexibility of the model to meet the characteristics and needs of individual programs.

Process indicators of fidelity. Qualitative interviews with program staff and consultants were another, critical source of information about process indicators of fidelity. These data confirmed and added to the analysis of the consultant logs. Both sources of data showed that the main themes of consultant content were reflective practice, working with children and families, and work relationships. The interview data underscored the ways in which consultants adapted their work to fit the needs of the individual programs. While consultants spoke favorably of their training in "Diversity-Informed Tenets for Work with Infants, Children, and Families" (Harris Foundation 2016; Tenets Initiative, 2018), issues of diversity, equity, and inclusion were not a primary topic of consultation in most programs, reflecting an area for future growth in implementing IECMHC. The qualitative data also indicated similarities in the overarching needs of center-based early childhood and home visiting providers and how the Illinois Model can effectively support both types of programs.

Factors affecting implementation. As expected, it took time for consultants to build relationships with program supervisors and staff and develop processes for working together. At the same time, there were several factors that either eased implementation or made it more

challenging. These included the ease or difficulty of scheduling meetings with staff and supervisors; stability or instability of staff at all levels (director, supervision, and staff); and extent to which leaders and staff understood IECMHC and their readiness to engage with the consultant. Indeed, one of the primary facilitators in successfully implementing the model was strong leadership support for consultation. Despite challenges with scheduling and staff support, evaluation data on implementation dosage, adherence, and process indicate that the implementation of the Illinois Model was overall successful in both early childhood center-based programs and home visiting programs.

Research Question 2: What were the effects of the intervention on staff and supervisors? Were there differences between staff in programs receiving the intervention and those in comparison programs in measured outcomes?

The theory of change for the Illinois Model posits that mental health consultation will result in several changes in staff and supervisor reflective capacity, functioning, and relationships. We found positive changes on two standardized measures of staff reflective capacity and a relationship between increased reflective capacity and decreased burnout in a subsample of staff. However, we did not see changes in standardized measures of burnout or depression (assessed low at baseline) or in measures of staff–supervisor relationships (assessed relatively favorably at baseline). According to the quantitative data, it appeared that other factors—specifically, teacher position and race/ethnicity—had stronger effect on some outcomes than the intervention did.

At the same time, there was evidence of an intervention effect on teachers' and home visitors' practices. Interview data confirmed this and revealed the following shifts in practice: 1) active listening and deeper exploration of issues, 2) the ability to think critically about one's reactions and biases, 3) the ability to consider others' perspectives, and 4) the ability to establish or improve boundaries and be mindful of self-care.

Reflective capacity. Strengthening staff reflective capacity through reflective consultation is an important component of the Illinois Model of IECMHC. The growth in staff reflective capacity was evident in both quantitative and qualitative data, whereas changes in supervisors were only apparent in the analysis of qualitative data, likely because of a small sample. The intervention demonstrated positive effects on two measures of staff reflective capacity.

Being in the intervention group also significantly predicted lower emotional exhaustion at Time 3, a component of burnout as measured by the MBI, for a subsample of staff, which was similar demographically to the larger sample. Growth in reflective process and collaboration (measured by the PRPAS) also predicted lower levels of emotional exhaustion, but the intervention was a stronger predictor. Thus, receiving the Illinois Model and building reflective capacity could

mitigate staff burnout. However, we need additional research to understand better how these factors impact burnout.

We did not find significant differences between the groups over time on burnout, but we did find group differences in burnout by race and ethnicity. In particular, staff who identified themselves as White reported higher emotional exhaustion compared to all other racial and ethnic groups. Previous research has found that white providers tend to report higher burnout than Black and Hispanic providers (Salyers & Bond, 2001, in caseworkers; Garcia et al., 2020, in physicians). Although the reasons for these differences are unclear, it might reflect differences in perceived burnout or differences in willingness to acknowledge feelings of burnout.

We also found that teacher role had an effect on their views of supervision and relationships with supervisors. Lead teachers who received the intervention perceived their supervisor's fidelity and delivery quality (RSRS), efforts to build a bond or relationship with them (SWAI Rapport), and efforts toward specific goals and tasks expected to benefit clients (SWAI Client focused) to be poorer than lead teachers in the comparison programs. It is unclear what may have contributed to lead teachers in the intervention group having a more negative perception of the supervision they received compared to the comparison group. One possible explanation is that the lead teachers in the intervention group experienced reflective conversations with the consultant, which led them to realize that the supervision they received was not as reflective.

Teacher reflective capacity and child outcomes. We found interesting relationships between staff reflective capacity and child outcomes. Compared to teachers with lower reflective capacity, teachers with higher reflective capacity reported less teacher stress associated with children's behaviors, rated children's social and emotional strengths related to resilience greater, and rated children as having fewer problems with aggression, withdrawal, attention, and the control of intense emotions. Although directionality cannot be determined from these findings, strengthening reflective capacity might lead to lower teacher stress and may shift teachers' perceptions of children to be more positive and strengths based. Previous research has found an association between teacher stress and teacher perception of child behavior as negative (e.g., Friedman-Krauss et al., 2014). It is also possible that teachers' more positive views of children lead to less stress and greater reflective capacity, as stress limits one's ability to be reflective (Ferguson, 2018). As Roffey (2012) noted, "How teachers feel makes a difference to their ability to respond effectively to the challenges they face" (p. 8).

Teacher depression and child outcomes. It is not clear whether IECMHC can affect measured depression in staff in the same way it can affect reflective capacity. However, depression is a variable that has been included in prior research on IECMHC (Silver & Zinsser, 2020), and we included it in the logic model as a distal outcome. Although greater reflective capacity was associated with teachers perceiving child behavior more positively, a two-item depression

screen predicted more negative views of child behaviors. Teachers' depressive symptoms were associated with their perceptions of children as having more internalizing and externalizing behavioral concerns and less ability to self-regulate their emotions and behavior. This association has a few possible explanations, as we cannot attribute causality: teacher depression could lead teachers to perceive child behavior more negatively; teacher depression could result in children exhibiting more behavioral concerns; or children's behavioral concerns and poor self-regulation skills could exacerbate teacher depression. Additional research could help to clarify this relationship.

Classroom climate. Observations at the center-based programs showed that teachers in the intervention group were more successful at managing children's behavior by enforcing clear, consistent, and developmentally appropriate rules of behavior and using proactive and positive behavior strategies over time than teachers in the comparison group. Teachers in the intervention group were also more likely to promote holistic development through a child-centered and individualized approach over time than teachers in the comparison group, although this finding was a trend that did not reach statistical significance. These findings from the classroom observations suggest that center-based early childhood programs that received the intervention had a climate that better promoted mental health, particularly by responding to children in more positive, developmentally appropriate ways, than programs who did not receive the intervention.

Equity in classrooms. Moreover, greater equity was observed in the classrooms of programs that were receiving the intervention (CHILD Equity auxiliary dimension) than comparison programs. Diversity, equity, and inclusion is foundational to the Illinois Model. One core competency of the model is the consultant's ability to work effectively in diverse cultures and communities through cultural humility. These concepts were emphasized in consultant training before the initiative started and during the implementation of the model through ongoing training, supervision, and reflective learning opportunities. Consultants and program supervisors also had opportunities to attend workshops on the Diversity-Informed Tenets (Tenets Initiative, 2018). Thus, the finding that classrooms in the intervention group had higher ratings on the CHILD Equity auxiliary dimension is promising. However, it also underscores the need for further research on how the DEI core competency is reflected in home visiting and classroom practices and how to develop that competency.

Home visitor engagement. In the home visiting programs, we observed differences in the video-recorded observations of visits with staff who did and did not receive the intervention. Home visitors in the intervention group more frequently engaged in responsive behaviors during the home visit and elicited input on the content and activities of the home visit from parents than home visitors in the comparison group. In addition, there was a trend for home visitors in the intervention group to facilitate positive parent-child interactions and encourage

the parent's leadership in the visit more often than home visitors in the comparison group. When we analyzed the home visit observation items that most aligned with the Illinois Model—essentially creating an IECMHC scale using the Home Visit Rating Scales-Adapted & Extended (HOVRS-A+; Roggman et al., 2010)—we found that home visitors who received the intervention significantly increased on this scale over time at a greater rate than those in the comparison group.

Research Question 3: What were the potential effects of the intervention on parent and child well-being and behavior? Were there differences between parents and children in programs receiving the intervention and those in comparison programs?

Child behavior. The evaluation did not assess children's behavior directly. Instead, it relied on teachers' ratings. When teachers rated the severity of problems in children who they perceived to have behavioral problems on the Strengths & Difficulties Questionnaire, teachers in the intervention group reported less severe behavioral problems over time than teachers in the comparison group. This finding did not quite reach statistical significance but suggests that the intervention may have improved teachers' understanding and perception of child behavior. Along with the classroom observation findings, this result supports the theory of change that mental health consultation for teachers can change both their practices to be more supportive of children's social and emotional development and their perceptions of children's behavior. Because our measures were all teacher-reported, however, it is unclear whether these changes reflect actual change in children's behavior.

Contrary to some of the findings in the literature (e.g., Gilliam et al., 2016b), there were no racial or ethnic differences in teachers' assessments of children's behavior. However, consistent with the literature (e.g., LeBuffe & Naglieri, 2012), there were a number of significant findings related to child gender. Notably, teachers rated girls significantly differently than boys on most of the child assessment measures: the Preschool Expulsion Risk Measure, the Devereux Early Childhood Assessment measure of children's social-emotional behavior, and the Strengths & Difficulties Questionnaire. The findings suggest that gender was the strongest influence on teachers' perceptions of children's behavior—stronger than race and stronger than the effect of mental health consultation. It may also suggest that another area of focus for mental health consultation is helping teachers better understand gender differences in children's development and behavior.

Family-level home visiting outcomes. Parents whose home visitors were receiving the intervention tended to report higher satisfaction in their role as parents than parents whose home visitors were in the comparison group. Along with the findings from the home visit observations, which showed that the home visitors who received the Illinois Model were

generally more responsive to families and more supportive of parent-child interactions, compared with home visitors who did not receive the intervention, parents participating in visits with those home visitors also felt better about being a parent.

Features of the home visit were also associated with family-level outcomes. The home visitors' responsiveness to the family during home visits was associated with the parent's role satisfaction and parental report of a positive home environment. Home visitor practice aiming to facilitate parent-child interactions was associated with the parent's report of the responsiveness and positivity in their interactions with their child, suggesting that home visitor practices that are intended to support parent-child interactions do, in fact, strengthen parent-child interactions. Home visitor behaviors and aspects of the home visit predicted to be affected based on the IECMHC theory of change for were also associated with parental report of positive interactions with their children. These findings support the connection between home visitor practice (and the content of home visits) and family outcomes.

Other outcomes of interest. While the evaluation found promising results on a number of relevant outcomes, we found no differences in changes over time between the intervention and comparison groups on other standardized measures: supervisor-staff relationships, staff burnout, self-efficacy, and depression. Many factors may affect these outcomes but two worth noting are: (1) baseline scores tended to be low for depression and burnout measures and high for the relationship and self-efficacy measures, so there was less room for change and (2) over a third of staff in the comparison group had opportunities to consult with a mental health consultant during the initiative, albeit with models that differed from the Illinois Model.

Study Strengths and Contributions

Our study makes important contributions to the growing body of IECMHC research literature. Several areas merit mention here.

Comprehensive, cross-system field study of both implementation and outcomes. This evaluation was the first to study the implementation of a new model of IECMHC in multiple early childhood systems, both school-based and community-based, using a matched-comparison group design. Although the variability in participating programs and consultants posed challenges for implementation, data collection, and analysis, it reflected the goal and commitment of the Leadership Team to examine implementation in the diverse communities and programs characteristic of Illinois.

IECMHC in home visiting. Although there is a growing body of research on IECMHC, few studies have included home visiting. The Illinois Model prioritized access to consultation across multiple early childhood systems, including home visiting, which provided an opportunity to examine its effects in that setting. We included six home visiting programs in the study and collected data from home visiting program supervisors, home visitors, and families, including

recorded observations of home visits. We found positive effects of the intervention on home visitor practices as measured through the HOVRS-A+. Specifically, home visitors who received IECMHC were more responsive to families and prioritized facilitating parent-child interactions during home visits.

Reflective functioning. We examined whether reflective functioning mediated or explained the relationship between IECMHC and other provider/staff and child outcomes. While improvements in staff reflective functioning did predict lower levels of staff burnout, reflective functioning was not a mediator between the intervention and burnout. The intervention showed effects on staff burnout while controlling for reflective functioning. We also tested whether staff reflective functioning mediated the intervention effects on child outcomes. We found that it did not.

Innovative measures. Many of the tools we used in this study were developed recently to measure constructs that are outcomes central to IECMHC but are also difficult to measure. These constructs include reflective capacity and reflective supervision. First, to measure reflective capacity, we used the PRPAS (Heller, 2017). One limitation of the tool is that the Multiple Perspectives scale contains only one item and the research team had difficulty scoring this scale reliably. Although more research is needed to validate the tool, the PRPAS shows promise as a measure of change in reflective capacity. Second, we administered a standardized scale in the surveys to measure reflective capacity, the Reflective Functioning Questionnaire (Fonagy et al., 2016). Third, we used the Reflective Supervision Rating Scale (Ash, 2010), a measure developed to assess the content and structure of reflective supervision from the supervisee's perspective.

For the classroom observations, we used the Climate of Healthy Interactions for Learning and Development (CHILD; Gilliam & Reyes, 2017), a comprehensive observational assessment of the mental health climate of early care and education settings. The domains on the CHILD align very well with the aims and anticipated outcomes of IECMHC. Finally, based on the theory of change for the Illinois Model of IECMHC, the research team selected 13 items from the HOVRS-A+ (Roggman et al., 2010) and created a new IECMHC scale within the HOVRS-A+ for home visit observations.

Analytic approach. The samples were clustered at different levels (e.g., children within classrooms within programs), so many of our data points were not independent. We used linear mixed modeling (LMM) to account for the nested longitudinal data, missing values, and the many covariates. There were different numbers of staff and children/families per program (an unbalanced clustered dataset), and the amount of time between data points was important to include, both of which LMM can address. Previous IECMHC evaluations that used a matched-comparison group design used techniques such as analysis of variance/covariance and did not account for the clustered levels of the data (Conners-Burrow et al., 2012; Egeren et al., 2011;

Gilliam, 2014). This is the first IECMHC evaluation to use both a matched-comparison group design and multilevel modeling.

Study Limitations and Challenges

As with every research study, our evaluation had some limitations. We outline the major limitations below and explain how we think these affected the study and its findings. It is our hope that future research on IECMHC initiatives will take these issues into consideration during the planning phase to ensure the strongest possible research designs.

Study timeline. We conducted the implementation study simultaneously with the outcome study. Ideally, an implementation evaluation is conducted first to measure how the intervention is implemented and identify any barriers to implementation. An outcome study would occur only after there was evidence that the intervention or program was implemented as planned. This sequence would result in greater confidence that any observed outcomes could be attributed to a fully functional intervention, and any outcomes that were not observed were not due to implementation issues. However, we designed the evaluation to be responsive to the multiple information needs of the Leadership Team, making clear that the study would prioritize implementation and staff-level outcomes, while also examining its potential to affect children and families.

There were not enough eligible programs in each setting, region, and community type to conduct a randomized control trial of the Illinois Model, which is typically considered the "gold standard" in evaluation design. We were able to use a matched-comparison group design to allow us to measure change that could be attributed to the intervention. Experts still consider the matched comparison group design to be a rigorous design in situations where it is not possible to randomly assign participants to study groups (see, for example, Hanita, et al., 2017). However, a limitation of this design is that we could not match programs on all potentially relevant program and staff characteristics before implementation started. Although there were few significant differences between the intervention and comparison groups, they differed in staff education.

Comparison programs receiving consultation. The programs in our comparison group were functioning as "business as usual," which means that they continued program operations as normal during the study. Several programs were receiving, or had access to, mental health consultation during the study period. Although the consultation models in the comparison sites were different than what the intervention programs were receiving, this may have masked measurable change of the Illinois Model on the intervention group in our analyses. We also lacked comprehensive information about the form and content of consultation in the comparison group, which limits our ability to explain differences or lack of differences in some of our outcome measures.

Variability in consultant relationships with programs. There was considerable variability in the intervention programs' relationships with their mental health consultants. Some had an existing relationship with their consultant prior to the start of this study. Some had received consultation previously and were familiar with one another. Others had never had a consultant before and had to develop relationships. These variations provided an opportunity to observe how the Illinois Model will work once it is implemented more broadly On the other hand, as a test of the model, the evaluation had to try to take into account the variable lengths of time it took to build trust between the consultant and the staff across the programs and fully implement the model.

Measure limitations. As noted above, we selected a number of outcome measures developed over the past decade for use in evaluations of mental health consultation and related interventions. Although some measures have been used in diverse populations, one limitation is that others are still being tested and validated and may evolve further. Some measures do not have published psychometrics, and some also might not have been sensitive enough to detect changes in staff and supervisor well-being and relationships to show changes that occurred as a result of the Illinois Model of IECMHC. For example, most staff reported low levels of burnout, and fairly positive relationships with supervisors, which meant that there was not a lot of room for improvement over time. Other researchers have suggested that baseline ratings may be artificially inflated, limiting ability to measure progress. For example, Heller and colleagues (2011) suggested that asking teachers to report on their own growth retrospectively after engaging in IECMHC may be more valid for measures like the TOS than collecting self-report data at baseline.

Data collection. We collected data over three academic years, which caused some difficulty in consistent data collection processes. Children moved classrooms and left programs. During the study, some programs were closed or had reduced programming during the summer. Additionally, there was higher than expected turnover in program staff, including supervisors, as well as children in early childhood center-based programs and families in home visiting programs. In particular, our sample of program supervisors was smaller than ideal, given how important the consultant-supervisor relationships are to the intervention. Additionally, some programs do not provide email addresses for staff, requiring us to send the study survey to their personal email addresses.

One final issue with data collection was the quality of the consultant logs. These logs were used to track consultants' activities, time spent on each, and the individuals involved in the activity. Understanding the proportion of time spent in each activity allows evaluators to analyze the intervention's implementation, as well as outcomes associated with the activities. However, the quality of the consultant log data varied, as did the level of detail provided. In addition, the format of the database used to collect the consultant logs was changed midway through the

evaluation, resulting in a learning curve for the consultants and technological issues to troubleshoot, as well as missing data that the evaluation team had to collect from consultants directly. As other evaluations of IECMHC have found (e.g., Egeren et al., 2011), the consultant logs were not a completely accurate representation of implementation. However, the consultant logs provided substantial information on their activities with staff and the number of hours delivered to each program, which was important information for evaluating the implementation.

Child assessments. Unfortunately, we could not conduct the child assessments on a random sample of children. Instead, in order to manage the data collection, we asked teachers to select no more than eight children in their classroom whose parents had provided informed consent and who were likely to remain in the program the following year. Nonetheless, because the baseline data collection period was in the spring, transitions in staff and children during the summer resulted in a smaller sample of children who remained with the same teacher in the fall when the second data collection occurred.

Expulsion and suspension data. We did not measure suspensions or expulsions directly via program records. This lack of data severely limits our ability to delve into whether this intervention impacted these often inequitable disciplinary actions. We did ask program directors during interviews about excluding children, but almost all explained that it was against agency or program policy to suspend or expel children.[35]

Sample sizes and attrition. The participating programs experienced a higher than expected level of turnover in program staff and leaders during the implementation period. Turnover negatively affected both program implementation and the evaluation. In two or three programs, program directors who were eager to participate in the study left their positions and were replaced by new directors who did not understand the value of IECMHC. A few programs in the study were notified that they would not receive funding from a state agency, and one then laid off all of its staff. When they subsequently received funding, the programs had to hire new staff, and the new funding brought different staffing requirements. This prevented programs from rehiring staff who were laid off because they did not meet the new hiring requirements. The timing of this was also unfortunate, as it occurred after Time 2 data collection and new participants could no longer be enrolled in the study. Thus, although the response rate of 64% was similar to the rate we expected at baseline, attrition led to final sample sizes that were smaller than anticipated.

[35] Head Start Program Performance Standards (Standards (U.S. Department of Health and Human Services, 2016) prohibit programs from expelling or un-enrolling children from Head Start programs because of a child's behavior. In addition, the IL legislation (Public Act 100-0105) went into effect just before the study began, so all of the early childhood care and education programs in our study would have been prohibited from expelling children as of January 1, 2018.

Effect sizes for the kinds of outcomes targeted and measured in IECMHC tend to be small, and these data are clustered in levels, requiring large sample sizes to have the power to detect these effects. Although we conducted power analyses to try to determine the necessary sample sizes for a matched comparison longitudinal design, we were limited by the lack of comparable studies in the literature that reported findings on staff and child outcomes in programmatic consultation using matched comparison group design. Another limitation was the number of staff, supervisors, children, and families available at the participating programs. Additional program sites in the evaluation would have resulted in larger samples of supervisors, children, and families.

Implications and Recommendations

This pilot study demonstrated a number of strengths of the Illinois Model. Establishing relationships and promoting infant and early childhood mental health through the parallel process (Johnston & Brinamen, 2006, 2012) are the foundation of the model. The model is preventive, aiming to support the well-being of children and families by building the capacity of the adults who care for and work with children, rather than only responding when challenges arise. The model uses reflective practice and a social justice framework to support and strengthen the early childhood care and education workforce. Its flexibility allows the approach to be implemented into different programs in different early childhood settings, each with its own set of challenges and needs. The study also resulted in several important findings relevant to practitioners, policymakers, and researchers interested in understanding what IECMHC can accomplish for program staff, families, and children. In the section below, we highlight some important considerations and implications of this research.

Practice Implications: The Illinois Model

Mental health professionals successfully implemented the Illinois Model in diverse settings, ranging from community-based childcare to school-based pre-K to home visiting programs. The consultants were well-trained and supported throughout the implementation, but they also varied in experience, understanding of the model, and prior relationships with the participating programs. Given all these variations, the model seems to have the right balance of structure and flexibility to be used in various settings by well-supported consultants from varied backgrounds. Implementation was facilitated by the infrastructure that was established by the Mental Health Consultation Initiative, which encompassed more than this pilot study. Notably, the initiative has created a strong workforce development plan, started the development of a centralized data system, and obtained funding to continue to coordinate efforts to advance IECMHC across multiple early childhood systems.

Based on the results of the pilot study, our recommendations for the Illinois Model and its implementation fall into three main areas—program commitment to and readiness for implementation; flexibility of model; and workforce development, as follows.

Program Readiness and Commitment

- Ensure readiness of program staff to engage with consultant and establish structures for implementation. Complete a thorough readiness assessment prior to implementation to ensure all staff, not just directors and supervisors, understand the structure and process of the Illinois Model and are engaged from the beginning. Depending on their understanding, some programs might need more support to become ready to engage with the consultant. Most importantly, program leaders need to commit to make time for consultation, just as they do for supervision.
- Establish minimum requirements and clear expectations for the consultation, including a regular schedule of meetings and space for the consultant.
- Continue to monitor implementation through data collection and periodic check-ins to make sure structures and schedules are working. Provide booster trainings every six months for staff and leadership in the model's approach or more often during times of staff transition.

Model Flexibility

- Maintain the flexibility of the Illinois Model's approach. Given that program administrators and staff will have varying levels of readiness, some may need more support than others to fully understand and engage with a consultant. Program structure, size, and staff needs will affect the monthly amount of consultation required. Our study findings suggest that 10-12 hours per month is appropriate for many larger programs, for example, those with 10 or more staff. However, smaller programs that do not have the schedules to allow for regular reflective supervision sessions may not have this much time. A consistent structure and schedule, based on staff size, might be more important than a specified number of hours. In addition, given the time it took some consultants to establish relationships with program staff at the beginning, more hours in the early months might help to solidify these relationships and ensure that staff and supervisors understand the Illinois Model's approach to consultation.
- Maintain the current practices recommended by the Illinois Model while also being flexible in their implementation. For example, the model advocates that consultants meet with staff and their supervisors together rather than individually. This helps to ensure good communication and relationships between supervisors and staff. Although some study participants, including a few consultants, resisted this idea at the beginning, over time they came to understand its value. Yet, some programs found it very difficult to coordinate schedules and put it into practice.
- Explore and be open to other means of communication with administrators and staff. The unfortunate arrival of the COVID-19 pandemic as the pilot was wrapping up forced

some early childhood programs to experiment with the delivery of consultation services through virtual means.

Workforce Development

- Continue to monitor implementation with online data collection by consultants. Periodically share data with programs leaders and staff to help them understand the process and progress of regular consultation.
- Maintain ongoing supports and training for consultants. All consultants participating in this study appreciated the regular monthly supervision and ongoing opportunities to reflect and learn provided to them during the implementation. It is not clear whether the same intensity of support will be available as the model is disseminated across the state, but consultants clearly valued and improved their skills with these supports. These supports were particularly important for less experienced consultants, with more seasoned consultants serving as mentors for new consultants.
- Relatedly, try to match consultants and programs so that consultants have experience with the system in which they are working. We found that staff and supervisors appreciated consultants who understood the content, funding requirements, and structures of the program they were serving.
- Maintain the requirement that consultants have supervision and opportunities to participate in reflective learning groups. Although our study did not focus on issues of diversity, equity, and inclusion (DEI), it was clear that this area was difficult for many consultants in the study to address. Given the promise of IECMHC to affect these areas, consultants and program staff would all benefit from a deeper understanding of these issues at a system level as well as help in applying DEI principles in their work. It appears that regular opportunities to reflect with peers were very beneficial for a number of topics, particularly DEI.
- In addition, provide more in-depth training and support to help consultants implement the Diversity-Informed Tenets. This study found that although consultants were familiar with and endorsed the Tenets because of training, they varied in their knowledge of them and comfort in addressing them with program staff. Consultants also reported that it was challenging to find the appropriate time and space for sensitive and uncomfortable conversations about DEI, particularly when program leaders did not recognize the relevance of these issues. These findings suggest a need for more intense training and, perhaps, more effective strategies and tools for consultants to use in implementing the Tenets, including how to initiate conversations related to DEI with program staff and administrators in order to support their growth in being culturally sensitive, aware, and humble.

Policy Implications

Illinois Inclusion Policy and IECMHC

The Illinois preschool expulsion ban legislation (Public Act 100-0105) was passed just prior to the start of the study. This law prohibits any program receiving funding from ISBE or licensed by DCFS from expelling children for behavioral reasons as of January 1, 2018. IECMHC was highlighted as an important resource for staff in this legislation. If programs could no longer remove children, they would need alternative solutions and resources to support them. This study adds to the growing body of evidence suggesting IECMHC is an effective support for early childhood program administrators and staff to develop new strategies for working with children who they perceive as having difficulties managing their behaviors in the program setting.

Early Childhood Workforce

Although research shows that young children and families benefit from high-quality early childhood experiences, it is not easy for providers to achieve the level of quality necessary to support child development. The ability of early childhood center-based programs to meet the needs of children and their families depends, more than anything, on the professional development, knowledge, and skills of their staff. Over the past two decades, educational requirements for staff and program quality standards in publicly funded programs, including Head Start, state pre-K, and home visiting, have become increasingly rigorous (Bernoteit et al., 2016), and it has been difficult for the early childhood workforce to keep pace with new requirements. As a result, the workforce has widely varying qualifications, degrees, and credentials as well as compensation, which typically differ by funding stream. At the same time, several researchers report that the early childhood workforce experience high levels of job stress and burnout, which can lead to poorer classroom environments and higher rates of child suspension or expulsion and staff turnover (Davis & Perry, 2014; Gillaim 2005; Jennings, 2015; RAINE Group, 2014; Whitebook & Sakai, 2003).

Although IECMHC can support program staff facing these challenges, it is not enough to address all the current issues and inequities in early childhood systems. IECMHC cannot be well implemented in a fragile system or fix systemic issues that contribute to staff stress, burnout, and turnover. For example, in community-based programs in this study, insufficient staff prevented consultants from facilitating reflective supervision because the supervisor had to serve as backup for a staff member. How can the Illinois Model make room to support programs with these kinds of barriers so that there is space for consultation rather than it feeling like an additional task on the list? Consultants showed themselves to be creative and adept at finding times to meet with supervisors or staff, but it was not easy. For IECMHC to be successful, staff must have time and space free of other responsibilities to meet with the consultant.

Research Implications

This study contributes to a growing body of research that has demonstrated positive effects of IECMHC for staff and families. However, additional research is needed to determine whether the Illinois Model of IECMHC leads to reductions in disparities, as theorized, as well as longer-term outcomes such as staff professionalization, staff retention, and reductions in problem behaviors and harsh disciplinary practices. We highlight some of our suggestions for additional research below.

- Conduct a follow-up study of program participants in this pilot to understand the sustained effects of consultation and structures that may or may not have been established to keep consultation in place. The COVID-19 pandemic has changed service delivery, especially in school-based and home visiting programs. Evaluating the implementation and sustainability of IECMHC during challenging times like these is necessary. Early childhood programs likely need the support of mental health consultants now more than ever.
- Do more study of implementation, paying special attention to differences between programs in different early childhood systems to better understand adaptations that should be made for different program types and differences within the childcare or home visiting systems. The Leadership Team has discussed other systems, including public health and Early Intervention, which might benefit from receiving the Illinois Model. However, there has been little research on implementing mental health consultation in these systems.[36]
- Explore the role of supervisors in IECMHC implementation and outcomes. Supervisors are less likely to be a focus of research on IECMHC but are integral to supporting the efforts of consultants to improve the knowledge and skills of frontline staff. Based on interviews with supervisors and consultants in this study, there was clear benefit for supervisors. However, our sample was very small, and standardized measures did not find differences between the two groups of supervisors. Thus, we recommend more study of the role of supervisors in implementing IECMHC, the challenges they experience in their work, and the supports they need to ensure that they work more effectively with frontline staff.
- To better understand outcomes of the Illinois Model for children and families, conduct an experimental or quasi-experimental study of the model with a longer study timeline and larger sample of children. For example, conduct a study that follows different cohorts over time as they transition to kindergarten. The child and family outcome data suggest that the Illinois Model has the potential to affect children and families in the long run, but more rigorous, longitudinal studies are needed to understand its impacts.

[36] A small pilot study of the Illinois Model in four public health settings in Illinois is nearing completion but otherwise, we are not aware of other published research on IECMHC in public health settings.

Furthermore, future research should measure the rates of child expulsion and suspension at the program level, if possible, to determine any impacts IECMHC may have on preventing expulsions and suspensions of young children.

- Examine how mental health consultation can improve the equity of early childhood settings for diverse populations in terms of managing children's problem behaviors. Classroom observations and staff surveys in this study revealed some differences by staff race and ethnicity. For example, on the CHILD, White teachers had lower scores on staff-child interactions and equity in their classrooms. It would be helpful to conduct further analyses of data from this study and other studies using the CHILD to examine the effect of teacher-children racial concordance and discordance (i.e., same vs. different racial identity) on the classroom climate.
- Work with other researchers to develop more sensitive measures of the changes expected from IECMHC to more clearly assess the outcomes and mechanisms of change of consultation, including reflective practice, supervisor-staff relationships, staff well-being, and ability to promote children's and families' social and emotional growth. The measures of reflective capacity used in this study show considerable promise, although the PRPAS takes time to administer and analyze. Furthermore, psychometric evidence is needed for some measures (e.g., TOS, GAS, PRPAS) to ensure reliability and validity, especially for use in evaluations of IECMHC.

Conclusion

Given the variations in implementation and the size of the samples in this evaluation, we find the outcomes for staff, children, and families to be promising. At the same time, the extent of changes in some of the outcomes indicates that there is room for further growth in staff, for example, in their reflective capacity and the social-emotional climate in classrooms. In addition, we need more study of outcomes, especially for supervisors, children, and families. Sample sizes for these three groups were very small. We were impressed that any of the changes in child and family measures were significant or trending towards significance, given the fact that these are more distal outcomes than staff outcomes.

References

Aarons, G. A., Fettes, D. L., Flores Jr., L. E., & Sommerfeld, D. H. (2009). Evidence-based practice implementation and staff emotional exhaustion in children's services. *Behaviour Research and Therapy*, *47*(11), 954–960.

Albritton, K., Mathews, R. E., & Anhalt, K. (2019). Systematic review of early childhood mental health consultation: Implications for improving preschool discipline disproportionality. *Journal of Educational and Psychological Consultation, 29*(4), 444–472.

Alkon, A., Ramler, M., & MacLennan, K. (2003). Evaluation of mental health consultation in child care centers. *Early Childhood Education Journal*, *31*(2), 91–99.

Allen, M. D., & Green, B. L. (2012). A multilevel analysis of consultant attributes that contribute to effective mental health consultation services. *Infant Mental Health Journal*, *33*(3), 234–245.

Arnett, J. (1989). *Caregiver interaction scale*. Princeton, NJ: Educational Testing Service.

Ash, J. (2009). *Kid Connects implementation manual*. Unpublished document.

Ash, J. (2010). *Reflective Supervision rating scale*. Unpublished measure.

Balazs, C. L., & Morello-Frosch, R. (2013). The three Rs: How community-based participatory research strengthens the rigor, relevance, and reach of science. *Environmental Justice*, *6*, 9–16.

Bernoteit, S. A., Darragh Ernst, J. C., & Latham, Nancy I. (Eds). (2016). *Voices from the field: Collaborative innovations in early childhood educator preparation*. Edwardsville, IL: Illinois Education Research Council and Illinois Board of Higher Education.

Bernstein, B.J. & Edwards, R.C. (2012): Supporting early childhood practitioners through relationship-based, reflective supervision. *NHSA Dialog: A Research-to-Practice Journal for the Early Childhood Field*, 15:3, 286–301http://dx.doi.org/10.1080/15240754.2012.694495

Buettner, C. K., Jeon, L., Hur, E., & Garcia, R. E. (2016). Teachers' social–emotional capacity: Factors associated with teachers' responsiveness and professional commitment. *Early Education and Development*, *27*(7), 1018–1039.

Bleecker, T. & Sherwood, D.L. (2003). *A final report of the 2001–2002 evaluation of the high quality child care mental health consultation initiative.* San Francisco, CA: Department of Public Health, Community Health Programs.

Brennan, E. M., Bradley, J. R., Allen, M. D., & Perry, D. F. (2008). The evidence base for mental health consultation in early childhood settings: Research synthesis addressing staff and program outcomes. *Early Education and Development, 19*(6), 982–1022.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, J., & Balain, S. (2007). A conceptual framework for implementation fidelity. *Implementation Science, 2*(1), 40.

Center for Prevention Research and Development. (2011). *Early Childhood Mental Health Consultation Project—HFI/PTS: Draft Report*. Champaign, IL: CPRD, Institute of Government and Public Affairs, University of Illinois.

Cigala, A., Venturelli, E., & Bassetti, M. (2019). Reflective Practice: A Method to Improve Teachers' Well-Being. A Longitudinal Training in Early Childhood Education and Care Centers. *Frontiers in Psychology, 10,* 2574.

Cohen, E., & Kaufmann, R. K. (2000). *Early childhood mental health consultation*. Center for Mental Health Services, Substance Abuse and Mental Health Services Administration, US Department of Health and Human Services.

Colorado Department of Public Health & Environment, Early Childhood Evaluation Unit (2020). *Mental Health Consultation Evaluation: Summary of interviews with Home Visitors, Supervisors, & Mental Health Consultants*. https://www.iecmhc.org/documents/MHC%20in%20Home%20Visting%20Summary%20of%20Findings.pdf

Colvard, J. & Cohen, J. (2019) Cross-system collaboration to better support babies in Illinois: Developing a unified model for infant and early childhood mental health consultation. *Zero to Three,* 1-5. https://www.zerotothree.org/document/1167

Colwell, N., Gordon, R. A., Fujimoto, K., Kaestner, R., & Korenman, S. (2013). New evidence on the validity of the Arnett Caregiver Interaction Scale: Results from the early childhood longitudinal study-birth cohort. *Early Childhood Research Quarterly*, *28*(2), 218–233.

Conners-Burrow, N. A., Whiteside-Mansell, L., McKelvey, L. (2012). Improved classroom quality and child behavior in an Arkansas early childhood mental health consultation pilot project. *Infant Metal Health Journal, 33*(3), 256–264.

Cigala, A., Venturelli, E., & Bassetti, M. (2019). Reflective Practice: A method to improve teachers' well-being. A longitudinal training in early childhood education and care centers. *Frontiers in Psychology*, *10*.

Crusto, C. A., Whitson, M. L., Feinn, R., Gargiulo, J., Holt, C., Paulicin, B., Simmons, W., & Lowell, D. I. (2013). Evaluation of a mental health consultation intervention in preschool settings. *Best Practices in Mental Health: An International Journal, 9(2*), 1–21

Cucchi, A., Hampton, J. A., & Moulton-Perkins, A. (2018). Using the validated Reflective Functioning Questionnaire to investigate mentalizing in individuals presenting with eating disorders with and without self-harm. *PeerJ*, *6*, e5756.

Daro, D. (2010). Replicating evidence-based home visiting models: A framework for assessing fidelity. *Supporting Evidence-Based Home Visiting to Prevent Child Maltreatment*, *Brief 3*. Chicago, IL: Chapin Hall.

Daro, D., McCurdy, K., Falconnier, L., & Stojanovic, D. (2003). Sustaining new parents in home visitation services: Key participant and program factors. *Child Abuse & Neglect, 27*(10), 1101–1125.

Davis, A., & Perry, D. F. (2016). *Pennsylvania Early Childhood Mental Health Consultation Project: External evaluation report*. Washington, DC: Georgetown University Center for Child and Human Development. https://www.iecmhc.org/wp-content/uploads/2020/12/PA_ECMHC_Report_Updated_executive-summary.pdf

Davis, A., & Perry, D. F. (2014). *Healthy Futures: Year four evaluation of early childhood mental health consultation*. Washington, DC: Georgetown University Center for Child and Human Development.

Davis, A. E., Perry, D. F., & Rabinovitz, L. (2020). Expulsion prevention: Framework for the role of infant and early childhood mental health consultation in addressing implicit biases. *Infant Mental Health Journal, 41,* 327–339.

Dougherty, A. (2013). *Psychological consultation and collaboration in school and community settings*. Stamford, CT: Cengage Learning.

Dunst, C. J. (2015). Improving the design and implementation of in-service professional development in early childhood intervention. *Infants & Young Children, 28*(3), 210–219. doi:10.1097/iyc.0000000000000042

Duran, F., Hepburn, K., Irvine, M., Kaufmann, R., Anthony, B., & Horen, N. (2009). *What works? A study of effective childhood mental health consultation programs*. Georgetown Center for Child and Human Development. Available from www.gucchd.georgetown.edu/78358.html

Durlak, J. A. (2015). Studying program implementation is not easy but it is essential. *Prevention Science, 16*(8), 1123-1127.

Efstation, J. F., Patton, M. J., & Kardash, C. M. (1990). Measuring the working alliance in counselor supervision. *Journal of Counseling Psychology, 37*(3), 322.

Egger, H. L., & Angold, A. (2006). Common emotional and behavioral disorders in preschool children: Presentation, nosology, and epidemiology. *Journal of Child Psychology and Psychiatry, 47,* 313–337.

Egeren, L. A. V., Kirk, R. H., Brophy-Herb, H. E., Carlson, J. S., Tableman, B., & Bender, S. L. (2011). *An interdisciplinary evaluation report of Michigan's Childcare Expulsion Prevention (CCEP) initiative*. Lansing: Michigan State University.

Erchul, W. P., & Martens, B. K. (2002). *School consultation: Conceptual and empirical bases of practice (2nd ed.).* New York, NY US: Kluwer Academic/Plenum Publishers.

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.

Fixen, D. L., Blasé, K. A., Horner, G., & Sugai, G. (2009). *Scaling up evidence-based practices in education. Scaling-Up Brief.* Chapel Hill, NC: The University of North Carolina, FPG Child Development Institute (issue #1). Retrieved from https://eric.ed.gov/?id=ED507440

Fonagy, P., Luyten, P., Moulton-Perkins, A., Lee, Y. W., Warren, F., Howard, S., Ghinai, R., Fearon,

P, & Lowyck, B. (2016). Development and validation of a self-report measure of mentalizing: The reflective functioning questionnaire. *PLoS One*, *11*(7), e0158678.

Garcia, L. C., Shanafelt, T. D., West, C. P., Sinsky, C. A., Trockel, M. T., Nedelec, L., Maldonado, Y. A., Tutty, M., Dyrbye, L. N., & Fassiotto, M. (2020). Burnout, depression, career satisfaction, and work-life integration by physician race/rthnicity. *JAMA Network Open, 3*(8):e2012762. doi:10.1001/jamanetworkopen.2020.12762

Geller, S. & Lynch, K. (1999). Teacher Opinion Scale. Unpublished measure.

Gilkerson, L., Hofherr, J., Heffron, M. C., Sims, J. M., Jalowiec, B., Bromberg, S. R., & Paul, J. J. (2012). Implementing the Fussy Baby Network [R] Approach. *Zero to Three, 33*(2), 59–65.

Gill, S., Greenberg, M. T., Moon, C., & Margraf, P. (2007). Home visitor competence, burnout, support, and client engagement. *Journal of Human Behavior in the Social Environment, 15*(1), 23–44. doi:10.1300/j137v15n01_02

Gilliam, W. S. (2005). *Prekindergarteners left behind: Expulsion rates in state prekindergarten systems*. New York, NY: Foundation for Child Development. https://medicine.yale.edu/childstudy/zigler/publications/National%20Prek%20Study_expulsion%20brief_34775_5379_v1.pdf

Gilliam, W. S. (2007). *Early childhood consultation partnership: Results of a random-controlled evaluation* (Final report and executive summary). New Haven, CT: Child Study Center, Yale University School of Medicine.

Gilliam, W. S. (2008). *Development of the preschool mental health climate scale: Final report*. New Haven, CT: Yale Child Study Center (unpublished).

Gilliam, W. S. (2008). Implementing policies to reduce the likelihood of preschool expulsion. *Foundation for Child Development Policy Brief*, *7*, 2–10. http://medicine.yale.edu/childstudy/zigler/images/prekexpulsionbrief2_tcm350-34772.pdf

Gilliam, W. S. (2014). *Early childhood consultation partnership: Results across three statewide random-controlled evaluations.* New Haven, CT: Yale School of Medicine, Child Study Center.

Gilliam, W. S., Maupin, A., Reyes, C. (2016a). Early childhood mental health consultation: Results of a statewide random-controlled evaluation. *Journal of the American Academy of Child & Adolescent Psychiatry, 55*(9), 754–761. https://doi.org/10.1016/j.jaac.2016.06.006

Gilliam, W. S., Maupin, A., Reyes, C., Accavitti, M., & Shic, F. (2016b). Do early educators' implicit biases regarding sex and race relate to behavior expectations and recommendations of preschool expulsions and suspensions? Yale University Child Study Center. https://medicine.yale.edu/childstudy/zigler/publications/Preschool%20Implicit%20Bias%20Policy%20Brief_final_9_26_276766_5379_v1.pdf

Gilliam, W. S., & Reyes, C. R. (2017). *Climate of Healthy Interactions for Learning & Development, Draft Manual* [Unpublished manual]. Edward Zigler Center in Child Development & Social Policy, Yale Child Study Center.

Gilliam, W. S., & Reyes, C. R. (2018). Teacher decision-making factors that lead to preschool expulsion: Scale development and preliminary validation of the Preschool Expulsion Risk Measure. *Infants and Young Children*, *31*(2), 93–108.

Gilliam, W. S., & Shahar, G. (2006). Preschool and child care expulsion and suspension: Rates and predictors in one state. *Infants & Young Children*, *19*(3), 228–245.

Giordano, K. (2019). Suspension & Expulsion Practices in NJ Community Childcare Centers: An Executive Summary. Union, NJ: Kean University. doi: 10.13140/RG.2.2.21751.98724.

Giordano, K., Interra, V.L., Stillo, G.C., Mims, A.T., & Block-Lerner, J. (2020). Associations between child and administrator race and suspension and expulsion rates in community childcare programs. Early Childhood Education Journal. doi: 10.1007/s10643-020-01049-1

Goodman, R. (1997). The Strengths and Difficulties Questionnaire: a research note. *Journal of Child Psychology and Psychiatry, 38*(5), 581-586.

Goodson, B. D., Mackrain, M., Perry, D. F., O'Brien, K., & Gwaltney, M. K. (2013). Enhancing home visiting with mental health consultation. *Pediatrics*, 132, S180–S190.

Green, B. L., Everhart, M., Gordon, L., & Gettman, M. G. (2006). Characteristics of effective mental health consultation in early childhood settings: Multilevel analysis of a national survey. *Topics in Early Childhood Special Education, 26*, 142–152.

Green, B., Malsch, A., Kothari, B., Busse, J., & Brennan, E. (2012). An intervention to increase early childhood staff capacity for promoting children's social-emotional development in preschool settings. *Early Childhood Education Journal*, 40 (2), 123–132. doi:10.1007/s10643011-0497-2

Green, B. L., Simpson, J., Everhart, M. C., Vale, E., & Garcia Gettman, M. (2004). Understanding integrated mental health services in Head Start: Staff perspectives on mental health consultation. *NHSA Dialog: A Research-to-Practice Journal for the Early Intervention Field*, *7*(1), 35–60.

Hanita, M., Ansel, D., & Shakman, K. (2017). *Matched-Comparison Group Design: An Evaluation Brief for Educational Stakeholders.* Education Development Center. https://oese.ed.gov/files/2019/03/02-09-MatchedComparison.pdf

Hansen, W. B. (2014). Measuring fidelity. In *Defining prevention science* (pp. 335-359). Boston, MA: Springer.

Harris Foundation. (2016). Illinois Action Plan to Integrate Early Childhood Mental Health Consultation into Family- and Child-Serving Systems, Prenatal through Age 5. https://www.irvingharrisfdn.org/wp-content/uploads/2017/09/Exec-Summary-IL-Action-Plan.pdf

Harris, P. A., Taylor, R., Thielke, R., Payne, J., Gonzalez, N., & Conde, J. G. (2009). Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of biomedical informatics*, *42*(2), 377–381.

Harris, P. A., Taylor, R., Minor, B. L., Elliott, V., Fernandez, M., O'Neal, L., . . . REDCap Consortium. (2019). The REDCap consortium: Building an international community of software platform partners. *Journal of Biomedical Informatics*, *95*, 103208.

Heffron, M. C. (2005). Reflective supervision in infant, toddler, and preschool work. In K. M. Finello (Ed.), *The Handbook of Training and Practice in Infant and Preschool Mental Health* (pp. 114–136). John Wiley & Sons.

Heller, S. S., Boothe, A., Keyes, A., Nagle, G., Sidell, M., & Rice, J. (2011). Implementation of a mental health consultation model and its impact on early childhood teachers' efficacy and competence. *Infant Mental Health Journal, 32*(2), 143–164.

Heller, S. (2017). *Provider Reflective Practice Assessment Scales*, Unpublished measure. Tulane University.

Hennigan, L., Upshur, C. C., & Wenz-Gross, M. (2004). Together for Kids: Second Year Report: A Project of Community Healthlink, Inc. Implementation Science and Practice Advances Research Center Publications. https://escholarship.umassmed.edu/psych_cmhsr/555

Hepburn, K. S., Perry, D. F., Shivers, E. M., & Gilliam, W. S. (2013). Early childhood mental health consultation as an evidence-based practice. *Zero to Three, 33*(5), 10–19.

Illinois Early Childhood Asset Map. (2020). Community reports. https://iecam.illinois.edu/

Jacobs, S. R., Weiner, B. J., & Bunger, A. C. (2014). Context matters: measuring implementation climate among individuals and groups. *Implementation Science, 9*(1), 46.

Jennings, P. A. (2015). Early childhood teachers' well-being, mindfulness, and self-compassion in relation to classroom quality and attitudes towards challenging students. *Mindfulness, 6*(4), 732–743.

Jennings, P.A., Jeon, L., & Roberts, A.M. (2020) Introduction to the special issue on early care and education professionals' social and emotional well-being, *Early Education and Development, 31* (7), 933–939. Retrieved from https://doi.org/10.1080/10409289.2020.1809895

Jeon, L., Buettner, C. K., Grant, A. A., & Lang, S. N. (2019). Early childhood teachers' stress and children's social, emotional, and behavioral functioning. *Journal of Applied Developmental Psychology*, *61*, 21–32.

Johnston, K., & Brinamen, C. (2006). *Mental health consultation in child care: Transforming relationships among directors, staff, and families*. Washington, DC: Zero to Three Press.

Katch, L. & Burkhardt, T. (2021). *Development and validation of the Infant Crying and Parent Well-being screening tool*. Manuscript under review.

Kaufmann, R.K., Perry, D.F., Hepburn, K., & Duran, F. (2012). Assessing fidelity for early childhood mental health consultation: Lessons from the field and next steps. *Infant Mental Health Journal*, *33*(3), 274–282. doi:10.1002/imhj.21337

Korfmacher, J., Sparr, M., Chawla, N., Fulford, J., & Fleming, J. (2012). *Illinois Prevention Initiative (PI) Birth to Three Program Evaluation*. Chicago, IL: Herr Research for Children and Social Policy Erikson Institute.

Kroenke, K., Spitzer, R. L., & Williams, J. B. (2003). The Patient Health Questionnaire-2: Validity of a two-item depression screener. *Medical Care*, 1284-1292.

Krysik, J., & Lecroy, C. W. (2012). Development and initial validation of an outcome measure for home visitation: The healthy families parenting inventory. *Infant Mental Health Journal*, *33*(5), 496–505.

Langkamp, D. L. (2003). *Evaluation of the Early Childhood Mental Health Initiative.* Akron, OH: Ohio Department of Mental Health, Office of Program Evaluation and Research.

LeBuffe, P., & Naglieri, J. (2012). *Devereux Early Childhood Assessment for Preschoolers Second Edition: User's Guide and Technical Manual*. Villanova, PA: Devereux Center for Resilient Children.

LeCroy & Milligan Associates, Inc. (2004). *Healthy Families Parenting Inventory*.

Lee, E., Esaki, N., Kim, J., Greene, R., Kirkland, K., & Mitchell-Herzfeld, S. (2013). Organizational climate and burnout among home visitors: Testing mediating effects of empowerment. *Children and Youth Services Review, 35*(4), 594–602. doi:10.1016/j.childyouth.2013.01.011

Lowenstein, L. F. (1991). Teacher stress leading to burnout: Its prevention and cure. *Education Today, 41*(2), 12–16.

Mackrain, M. (2011). A day in the life of an early childhood mental health consultant. Center for Early Childhood Mental Health Consultation: Georgetown University Center for Child and Human Development. http://www.ecmhc.org/documents/Day_in_the_Life_MH_Consultant.pdf

Mackrain, M., Perry, D., Yates, D. & Pell, K. (2011). *Early childhood mental health consultation (ECMHC) guide: Standards, rationale, and guidance for the state of Maryland*. Maryland Department of Education. http://www.marylandpublicschools.org

Madill, R., Halle, T., Gebhart, T., & Shuey, E. (2018). *Supporting the psychological wellbeing of the early care and education workforce: Findings from the National Survey of Early Care and Education*. OPRE Report #2018-49. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Malik, R. (2017). New data reveal 250 preschoolers are suspended or expelled every day. Center on American Progress. https://www.americanprogress.org/issues/early-childhood/news/2017/11/06/442280/new-data-reveal-250-preschoolers-suspended-expelled-every-day/.

Manz, P. H., & Ventresco, N. E. (2019). Observing home-visiting quality across time: A longitudinal reliability study of the Home Visit Rating Scales. *Infant Mental Health Journal*, *40*(3), 363–379.

Maslach, C., Jackson, S. E., & Leiter, M. P. (1996). *MBI: Maslach burnout inventory*. Sunnyvale, CA: CPP, Incorporated.

Meek, S. E., & Gilliam, W. S. (2016). *Expulsion and suspension as matters of social justice and health equity*. Washington, DC: National Academy of Medicine. https://nam.edu/wp-content/uploads/2016/10/Expulsion-and-Suspension-in-Early-Education-asMatters-of-Social-Justice-and-Health-Equity.pdf

Michalopoulos, C., Faucetta, K., Hill, C. J., Portilla, X. A., Burrell, L., Lee, H., & Knox, V. (2019). *Impacts on family outcomes of evidence-based early childhood home visiting: Results from the mother and infant home visiting program evaluation.* OPRE Report 2019-07. Washington, DC: Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Morrison, M. O., & Bratton, S. C. (2010). Preliminary investigation of an early mental health intervention for Head Start programs: Effects of child teacher relationship training on children's behavior problems. *Psychology in the Schools, 47*(10), 1003–1017. doi:10.1002/pits.20520

Nadiv, S., (2013). *Early childhood mental health consultation outcomes monitoring system (OMS) manual*. The Institute for Innovation & Implementation University of Maryland, School of Social Work. https://theinstitute.umaryland.edu/topics/ebpp/ecmhc.cfm

Parlakian, R. (2001). *Look, listen, and learn: Reflective supervision and relationship-based work*. Zero to Three.

Patton, M. Q. (2010). *Developmental Evaluation*. Guilford Press.

Patton, M. Q. (2012). *Essentials of utilization-focused evaluation*. Sage.

Perry, D. F. (2013). *Healthy Futures: Year Three Evaluation of Early Childhood Mental Health Consultation by the District of Columbia Department of Mental Health*. https://dbh.dc.gov/sites/default/files/dc/sites/dmh/publication/attachments/HealthyFuturesThreeYearEvaluationReport.pdf

Perry, D. F., Allen, M. D., Brennan, E. M., & Bradley, J. R. (2010). The evidence base for mental health consultation in early childhood settings: A research synthesis addressing children's behavioral outcomes. *Early Education and Development*, *21*(6), 795–824.

Perry, D. F. & Conners-Burrow, N. (2016). Addressing early adversity through mental health consultation in early childhood settings. *Family Relations, 65(1*), 24–36.

Perry, D. F. (2011). The evidence base for early childhood mental health consultation. USDHS Pub No. 90YD0268. Georgetown University Center for Child and Human Development. http://familymedicine.uams.edu

Perry, D. F., Dunne, M. C., McFadden, L., & Campbell, D. (2008). Reducing the risk for preschool expulsion: Mental health consultation for young children with challenging behaviors. *Journal of Child and Family Studies*, *17*(1), 44–54.

Perry, D. F., & Linas, K. (2012). Building the evidence base for early childhood mental health

consultation: Where we've been, where we are, and where we are going. *Infant Mental Health Journal, 33*(3), 223–225. doi:10.1002/imhj.2133

Pianta, R. (2001). *Student–Teacher Relationship Scale–Short Form*. Lutz, FL: Psychological Assessment Resources, Inc.

Pittenger, S., Van Dyk, T., Schreier, A., Haikalis, M., Floor, M. F., & Hansen, D. (2015). Expanding mental health consultation in Early Head Start: Recommendations for supporting home visitors in increasing parental engagement. *NHSA Dialog, 18*, 119–135. https://journals.uncc.edu/dialog/article/view/288/0

Poghosyan, L., Aiken, L. H., & Sloane, D. M. (2009). Factor structure of the Maslach burnout inventory: an analysis of data from large scale cross-sectional surveys of nurses from eight countries. *International Journal of Nursing Studies*, *46*(7), 894–902.

The RAINE Group. (2014). *Early childhood mental health consultation: Protects and maximizes our national investment in early care and education.* Southwest Human Development.

Raver, C. C., Jones, S. M., Li-Grining, C. P., Metzger, M., Champion, K. M., & Sardin, L. (2008). Improving preschool classroom processes: Preliminary findings from a randomized trial implemented in Head Start settings. *Early Childhood Research Quarterly, 23*(1), 10–26.

Raver, C. C., Jones, S. M., Li-Grining, C., Zhai, F., Metzger, M., Solomon, B. (2009). Targeting children's behavior problems in preschool classrooms: A cluster-randomized controlled trial. *Journal of Counseling and Clinical Psychology, 77*(2), 302–316.

Reyes, C. R. & Gilliam, W. S. (2018). *Climate of healthy interactions for learning & development implementation guide.* Edward Zigler Center in Child Development & Social Policy Yale Child Study Center.

Roffey, S. (2012). Pupil wellbeing—Teacher wellbeing: Two sides of the same coin? *Educational and Child Psychology, 29*(4), 8.

Roggman, L. A., Cook, G. A., Innocenti, M. S., Jump Norman, V. K., Christiansen, K., Boyce, L. K., Aikens, N., Boller, K., Paulsell, D., & Hallgren, K. (2010). Home Visit Rating Scales—Adapted and to Excellence (HOVRS-A+) v2.0. Unpublished measure.

Saldaña, J. (2015). *The coding manual for qualitative researchers.* London: Sage.

Salyers, M. P., & Bond, G. R. (2001). An exploratory analysis of racial factors in staff burnout among assertive community treatment workers. *Community Mental Health Journal*, *37*(5), 393–404.

Schodt, S., Parr, J., Araujo, M. C., & Rubio-Codina, M. (2015). *Measuring the quality of home-visiting services: a review of the literature*. Inter-American Development Bank.

Shahmoon-Shanok, R. (2006). Reflective supervision: A method of implementation. In G. M. Foley & J. D. Hochman (Eds.), *Mental health in early intervention: Achieving unity in principles and practice* (pp. 343–381). Baltimore, MD: Paul H. Brookes.

Shivers, E.M. (2011). *Smart Support Year 1 evaluation report: Arizona's early childhood mental health consultation program*. Prepared for Southwest Human Development, with support from First Things First.

Shivers, E. M. (2015). Arizona's Smart Start Support evaluation report: The first four years. Indigo Cultural Center for Southwest Human Development, with support from First Things First. https://www.swhd.org/wp-content/uploads/2015/02/Smart-Support_FNL_web.pdf

Shivers, E. M., Lowe-Fotos, A., Yazzie, D., & Lopez, D. (2018). *Advancing racial equity in early childhood education through mental health consultation.* Presentation at the Zero to Three Annual Conference, Denver, CO.

Shulman, C. (2016). *Research and Practice in Infant and Early Childhood Mental Health*, Volume 13 of the series *Children's Well-Being: Indicators and Research.* Springer.

Silver, H. C. & Zinsser, K. M. (2020). The interplay among early childhood teachers' social and emotional well-being, mental health consultation, and preschool expulsion. *Early Education and Development,* 1–18. https://doi.org/10.1080/10409289.2020.1785267

Smith, M. V., Gotman, N., Lin, H., & Yonkers, K. A. (2010). Do the PHQ-8 and the PHQ-2 accurately screen for depressive disorders in a sample of pregnant women? *General Hospital Psychiatry, 32*(5), 544–548.

Spielberger, J., Gouvêa, M., Winje, C., Burkhardt, T., Gitlow, E., & Pacheco-Applegate, A. (2019). *Evaluation of the Fussy Baby Network Condensed FAN Training: Final report.* Chicago, IL: Chapin Hall at the University of Chicago.

Stephan, S., Mettrick, J., Nadiv, S., Bedi, S., Chambers, K. L., Bruno, A., Perry, D. F., Holland, C. M., Darling-Kuria, N., & Pell, K. (2011). *Maryland's Early Childhood Mental Health Consultation Evaluation*. University of Maryland School of Medicine, Division of Child and Adolescent Psychiatry. http://hdl.handle.net/10713/7400

Susman-Stillman, A., Lim, S., Meuwissen, A., & Watson, C. (2020). Reflective supervision/consultation and early childhood professionals' well-being: A qualitative analysis of supervisors' perspectives. *Early Education and Development*, 1–18.

Tandon, S. D., Parillo, K. M., Jenkins, C., & Duggan, A. K. (2005). Formative evaluation of home visitors' role in addressing poor mental health, domestic violence, and substance abuse among low-income pregnant and parenting women. *Maternal and Child Health Journal*, *9*(3), 273–283.

Tenets Initiative (2018). Diversity-Informed Tenets for Work with Infants, Children & Families/Principios informados en la diversidad para trabajar con bebés, niños, niñas y familias. Chicago, IL: Irving Harris Foundation. https://diversityinformedtenets.org/download-the-tenets/

U.S. Census Bureau. (2020). *American Community Survey 2017: Demographic and housing estimates*, Peoria County, Illinois, and Cook County, Illinois. https://data.census.gov/cedsci/

U.S. Department of Education Office for Civil Rights (U.S. DOE OCR). (2016). 2013–2014 Civil rights data collection: A first look: Key data highlights on equity and opportunity gaps in our nation's public schools. https://www2.ed.gov/about/offices/list/ocr/docs/2013-14-first-look.pdf

U.S. Department of Health and Human Services. (2016). *Head Start Program Performance Standards.* U.S. Government Printing Office.

Upshur, C., Wenz-Gross, M., & Reed, G. (2009). A pilot study of early childhood mental health consultation for children with behavioral problems in preschool. *Early Childhood Research Quarterly*, *24*(1), 29–45.

Vogel, C. A., Caronongan, P., Thomas, J., Bandel, E., Xue, Y., Henke, J., Aikens, N., Boller, K., & Murphy, L. (2015). *Toddlers in Early Head Start: A Portrait of 2-Year-Olds, Their Families, and the Programs Serving Them. Volume 1: Age 2 Report. OPRE Report 2015-10.* Administration for Children & Families.

Vuyk, M. A., Sprague-Jones, J., & Reed, C. (2016). Early childhood mental health consultation: An evaluation of effectiveness in a rural community. *Infant Mental Health Journal*, 37(1), 66–79.

Watson, C., & Gatti, S. N. (2012). Professional development through reflective consultation in Early Intervention. *Infants & Young Children, 25*(2), 109–121. doi:10.1097/iyc.0b013e31824c0685

Webster-Stratton, C. (2005). *The Incredible Years: A training series for the prevention and treatment of conduct problems in young children*. In E. D. Hibbs & P. S. Jensen (Eds.), Psychosocial treatments for child and adolescent disorders: Empirically based strategies for clinical practice (pp. 507–555). American Psychological Association.

West, B. T., Welch, K. B., & Galecki, A. T. (2007). *Linear mixed models: a practical guide using statistical software.* London: Chapman & Hall/CRC.

Whitebook, M., & Sakai, L. (2003). Turnover begets turnover: An examination of job and occupational instability among child care center staff. *Early Childhood Research Quarterly, 18,* 273–293.

Zinsser, K. M., Zalauf, C. A., Das, N. V., & Silver, H. C. (2019). Utilizing social-emotional learning supports to address teacher stress and preschool expulsion. *Journal of Applied Developmental Psychology,61*, 33–42. https://doi.org/10.1016/j.appdev.2017.11.006

Appendix A

Figure A-1. Organizational Chart for the Infant/Early Childhood Mental Health Consultation Initiative

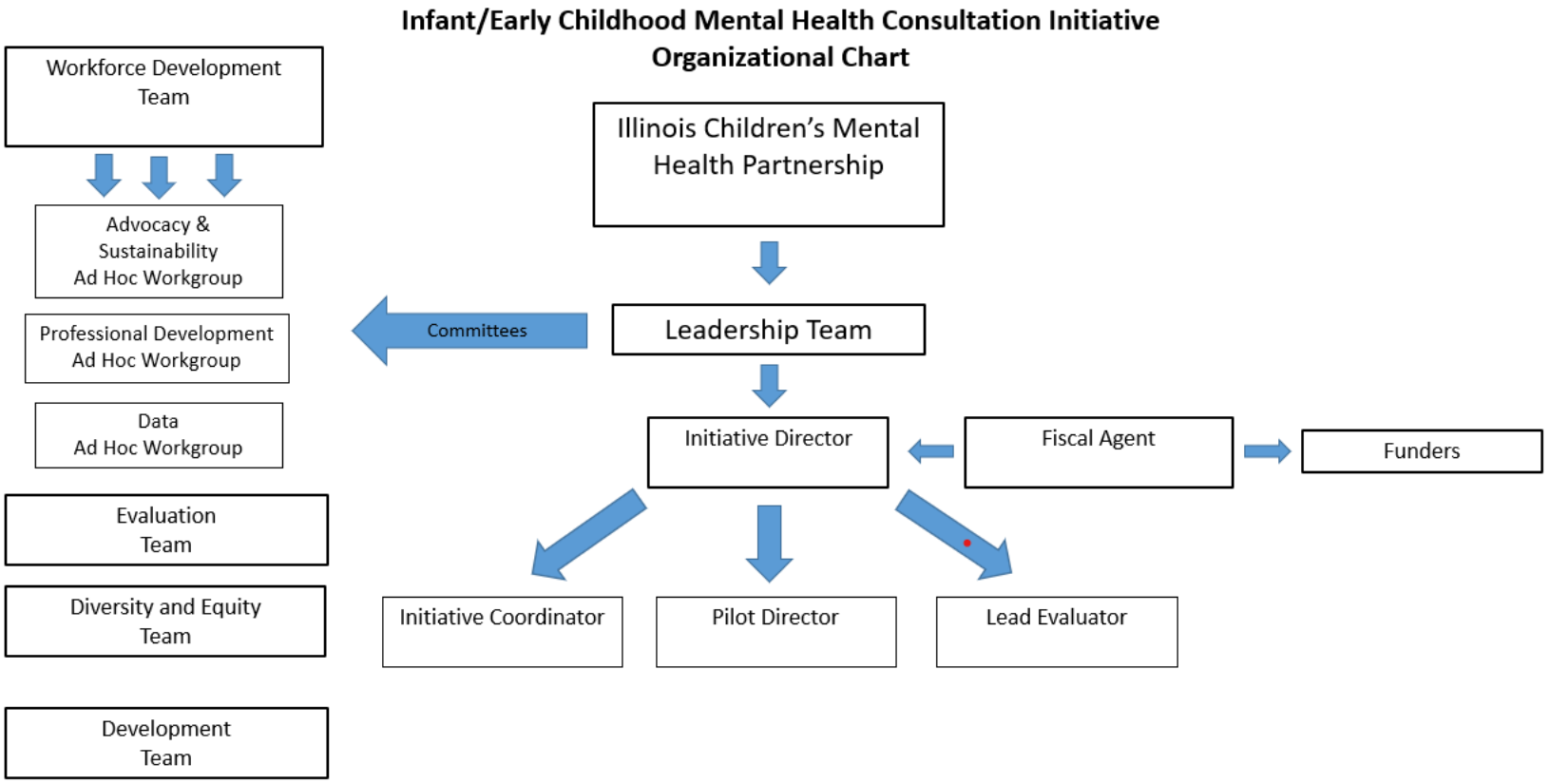


Table A-1. Baseline, Time 2, Time 3, and Time 4 Data Collection

Method	Source	Indicator	Baseline (#/%)	Time 2 (#/%)	Time 3 (#/%)	Time 4 (#/%)
Survey	Teacher/ Home Visitor	Sent	293	253	173	106
		Completed	180	164	130	86
		Percent	*61*	*65*	*75*	*83*
	Supervisor	Sent	31	29	21	16
		Completed	26	21	18	11
		Percent	*84*	*72*	*86*	*69*
	All Staff	Sent	324	282	194	122
		Completed	206	185	148	97
		Percent	*64*	*65*	***76***	*80*
Interview	Teacher/ Home Visitor	Requested	54	—	40	—
		Completed	41	—	28	—
		Percent	*76*	—	*70%*	—
	Supervisor/ Director	Requested	38	32	31	25
		Completed	30	30	24	22
		Percent	*79*	*94*	*77*	*88*
	All Staff	Requested	92	—	72	—
		Completed	71	—	52	—
		Percent	*77%*	—	***72%***	—
	Consultants		—	12 (100)	12 (100)	10 (100)
Classroom Observation	Teacher	Requested [a]	40	35	33	27
		Completed	40	35	33	27
		Percent	*100*	*100*	***100***	*100*
Child Assessments	Teacher	Requested	40	20	17	—
		Completed [b]	22	16	16	—
		Percent	*55*	*80*	***94***	—
		Total # children assessed	141	60	54	—
HV Video Recording	Video Recordings	Requested	26	14	12	12
		Completed	18	12	10	7
		Percent	***64***	***86***	***83***	***58***
	Post Video HV Survey	Requested	22	14	12	12
		Completed	18	12	10[b]	7
		Percent	*82*	*86*	*83*	*58*
	Post Video Parent Survey	Requested	24	14	12	12
		Completed	19	11	10	7
		Percent	*79*	*79*	*83*	*58*

Table A-2. Demographic Characteristics of Frontline Staff at Baseline[a,b]

	Full Baseline Sample		Analytic Sample[c]	
	Intervention	Comparison	Intervention	Comparison
Characteristic	*N* = 119	*N* = 76	*N* = 72	*N* = 64
Gender (%)	*N* = 119	*N* = 76	*N* = 72	*N* = 64
Female	97	100	96	100
Male	3	0	4	0
Race and Ethnicity (%)	*n* = 118	*n* = 75	*n* = 70	*n* = 63
Black	27	23	20	22
White	52	44	54	48
Latina/Hispanic	21	32	26	30
American Indian/ Native American	0	1	0	0
Age (%)	*n* = 118	*n* = 76	*n* = 71	*n* = 64
Under 20 years	0	4	0	5
20–29 years	27	20	24	17
30–39 years	26	30	30	30
40–49 years	27	20	28	17
50 or older	20	26	18	31
Education (%)[d]	*n* = 118	*n* = 75	*n* = 70	*n* = 62
Some college/no degree	19	23	14	24
Associate's Degree	24	35	20	32
Bachelor's Degree	49	27	57	27
Master's degree	9	16	9	16
Years of Experience	*n* = 105	*n* = 56	*n* = 64	*n* = 46
Mean (*SD*)	10.2 (8.84)	10.6 (7.74)	10.4 (9.66)	11.2 (8.21)
Range	1–43	1–39	1–43	1–39
Program Type	*N* = 119	*N* = 76	*N* = 72	*N* = 64
Early childhood center	83	93	83	94
Home visiting	17	7	17	6
Location (%)	*N* = 119	*N* = 76	*N* = 72	*N* = 64
Chicago area	50	57	49	53
Peoria area	50	43	51	47

[a]The full baseline sample includes staff from one EC center-based and one HV program that were not able to complete all data collection activities and, therefore, were not part of the analytic sample followed over time.
[b]Percentages reflect those who responded to the survey question.
[c]Analytic sample refers to the sample of staff who responded at baseline and at least one other time point and, therefore, could be included in analyses of change over time.
[d]The proportion of participants with each level of educational attainment was significantly different between the intervention and comparison groups in the analytic sample two groups ($\chi 2 = 11.84$, $p = .008$).

Table A-3. Demographic Characteristics of Program Supervisors at Baseline[a,b]

	Baseline Sample		Analytic Sample	
	Intervention	Comparison	Intervention	Comparison
Characteristic	*N* = 20	*N* = 7	*N* = 14	*N* = 5
Sex (%)	*N* = 20	*N* = 7	*N* = 14	*N* = 5
Female	85	100	86	100
Male	15	0	14	0
Race and Ethnicity (%)[b]	*n* = 18	*N* = 7	*n* = 12	*N* = 5
Black	28	43	17	40
White	44	29	50	40
Hispanic	17	29	25	20
Asian	6	0	8	0
Multiple races/ethnicities	6	0	0	0
Age (%)	*N* = 20	*N* = 7	*N* = 14	*N* = 5
30–39 years	25	43	29	40
40–49 years	55	14	57	20
50 or older	20	43	14	40
Education (%)	*N* = 20	*N* = 7	*N* = 14	*N* = 5
Associate's Degree	10	14	14	20
Bachelor's Degree	20	29	21	20
Master's Degree or above	70	57	65	60
Years of Experience	*n* = 17	*n* = 6	*n* = 11	*n* = 4
Mean (*SD*)	15.1 (9.01)	12.8 (11.75)	13.4 (6.27)	18.5 (10.08)
Range	2–32	1–32	5–26	8–32
Program type (%)	*N* = 20	*N* = 7	*N* = 14	*N* = 5
Early childhood center	75	71	79	80
Home visiting	25	29	21	20
Location (%)	*N* = 20	*N* = 7	*N* = 14	*N* = 5
Chicago area	50	71	57	60
Peoria area	50	29	43	40

[a] The full baseline sample includes supervisors from one EC center-based and one HV program that were not able to complete all data collection activities and, therefore, were not part of the analytic sample followed over time.
[b] Percentages reflect those who responded to the survey question. Two supervisors from implementation programs did not respond to the question about race/ethnicity.

Outcome Measures

The table below lists constructs and measures used in this study. In selecting them, we carefully reviewed measures used in previous evaluations of mental health consultation in early childhood settings and discussed them with the initiative's Evaluation Team. We prioritized measures with evidence of validity for the anticipated target population of the proposed evaluation. Although some of the measures are relatively new and have not been validated across diverse populations, they measure important constructs that are hypothesized to be outcomes of IECMHC.

Table A-4. Measures for IECMHC Evaluation: Evidence of Applicability in Early Childhood Settings and with Diverse Populations

Construct	Measures	Used in previous EC research	Used with diverse populations
Job burnout	Maslach Burnout Inventory	Hennigan et al., 2004; Langkamp, 2003	Hennigan et al., 2004; Poghosyan, Aiken, & Sloane, 2009
Reflective functioning	Reflective Functioning Questionnaire	N/A	Fonagy, Luyten, Moulton-Perkins, Lee, Warren, Howard et al., 2016; Cucchi, Hampton, & Moulton-Perkins, 2018
	Provider Reflective Practice Assessment Scales[a]	Validation of this measure is currently in process	Validation currently in process
Provider self-efficacy	Teacher Opinion Survey[a]	Duran et al., 2009; Egeren et al., 2011; Shivers, 2015; Stephan et al., 2011	Shivers, 2015[37]
Provider competence (skills and interactions)	Goal Achievement Scale[a]	Perry, 2013; Egeren et al., 2011	Perry, 2013
Consultant knowledge & skills	Knowledge & Skills Inventory	Erchul & Marten, 2002; Shivers, 2015	Erchul & Marten, 2002; Shivers, 2015
Child behavior & social-emotional development	Devereux Early Childhood Assessment for Preschoolers Second Edition (DECA-P2)	For review see, LeBuffe & Naglieri, 2012	For review see, LeBuffe & Naglieri, 2012
	Preschool Expulsion Risk Measure	Gilliam, 2014; Shivers, 2015	Shivers, 2015
	Strengths & Difficulties	Egeren et al., 2011; Perry, 2013	Child samples were not diverse

[37] Providers in Shivers (2015) were primarily White (54%) and Latino (30%), with Black providers under-represented (7%)

Construct	Measures	Used in previous EC research	Used with diverse populations
Parenting practices & parent-child interactions	Questionnaire Impact Supplement Healthy Families Parenting Inventory	Krysik & LeCroy, 2012	Krysik & LeCroy, 2012
Organizational/ Classroom Climate	Climate of Healthy Interactions for Learning and Development[1]	Newly developed (2016), the CHILD is a revision of the PMHCS, which has been widely used (e.g., Gilliam, 2014; Perry, 2013; Shivers, 2015; Stephan et al., 2011).	No published research yet, but all early childhood mental health consultants in CO, OH, MI, and NYC have been trained in the CHILD.
Home Visit Observations	Home Visit Rating Scales – Adapted and Extended	Korfmacher et al., 2012; Manz & Ventresco, 2019; Vogel et al., 2015	Schodt et al., 2015

[a] No published psychometrics.

The Provider Reflective Process Assessment Scales (PRPAS)

The Provider Reflective Process Assessment Scales (PRPAS; Heller, 2017) were administered to teachers and home visitors twice as part of their interviews, once at baseline and again one year later (Time 3). The PRPAS consists of six scales that correspond to different dimensions of reflective capacity. These are briefly summarized in Table A-5.

Table A-5. Provider Reflective Practice Assessment Scales (PRPAS)

Scale	Description of indicators within scale
Self-Knowledge	The extent to which the respondent (1) considers the impact of their own values, beliefs, and/or assumptions and how these may influence their words, actions, and thoughts; and (2) seeks to learn more about him/herself, sees themselves as a perpetual learner, and/or applies information learned.
Self-Regulation	The extent to which the respondent (1) takes time to pause and reflect before acting in emotionally charged situations, (2) does not rush to dismiss or repair client's uncomfortable (or negative) emotions and holds own uncomfortable emotions, and (3) recognizes the emotional climate of client and supports the client without adopting the client's emotions.
Multiple Perspectives	The extent to which the respondent is aware of the personal history, experiences, and culture of self and strives to understand those of the client and other important people in client's life and to help the client understand these differing perspectives and their impact on behavior.
Collaboration	The degree to which the respondent (1) approaches concerns from the perspective of inquiry (as opposed to inquisition) and explores potential solutions with the client; (2) does not respond to client out of an urge to fix, but rather slows down, develops a full understanding, and supports client in exploring potential solutions; and (3) is attuned to the potential impact of their words or actions on the client and takes time to contemplate how to approach a client in especially tenuous situations.
Process	The extent to which the respondent (1) recognizes that much of learning is experiential and occurs through relationships and (2) appreciates the complexity and richness of client's story and allows it to unfold.
Authentic Attitude	The extent to which the respondent (1) hears and responds to information from client or supervisor in an accepting, nonjudgmental, and supportive manner; (2) maintains a sense of wonder, interest, and compassion, and a sense of willingness to learn more; and (3) accepts clients for who they are and supports them without being judgmental, without letting preconceived ideas impact understanding, expectations, or support of them.

Note. Table above is used with permission from Sherry Heller, author.

In the administration of the instrument, teachers and home visitors were asked to talk about a parent or family with whom they had worked and found challenging with reference to the following topics: (1) their work with this family, (2) why the case was challenging, (3) how they went about addressing those challenges and working with the family, and (4) what they learned about themselves in their work with this family. They were asked to talk for 5 minutes, without interruption, to share their experiences, thoughts, and reflections. We were able to collect and

analyze baseline and Time 3 PRPAS data from 26 of the 58 teachers and home visitors in the focused sample.[38]

Analytic Approach: Linear Mixed Modeling

The first step is to select the model that runs and generate confidence intervals on the variance-covariance components. The methodology by West et al. (2007) recommends including all possible covariates and interaction terms with time in this step. However, it is very difficult to have a final model that includes all possible covariates for multiple reasons. Sometimes it is because two variables are highly correlated (i.e., multicollinearity) and cannot be included in the model together. Other times, there are variables that are not statistically significant and are dropped out of the model. We initially included all possible covariates and then removed variables that were not statistically significant or that were highly correlated with other variables. The model selected at this stage includes the variables that did not drop out and that generated confidence intervals on the variance-covariance components.

The following is the general specification of the LMM model:

$$Y_i = \underbrace{X_i\beta}_{\text{fixed}} + \underbrace{Z_i u_i + \varepsilon_i}_{\text{random}}$$

$$u_i \sim N(0, D)$$

$$\varepsilon_i \sim N(0, R_i)$$

Where u_i is the vector that contains the random effects of the model (site-level and staff-level variables). LMM estimates the random effects using a very flexible representation of D (the variance-covariance matrix (elements along the main diagonal of the D matrix represent the variances of each random effect in u_i, and the off-diagonal elements represent the covariances between two corresponding random effects). Thus, we had to estimate that matrix in order to get the random effects of the model and in order to get the complete results (fixed and random components) of this model. If we select a model with variables that do not allow this to happen then the results are not correct and the estimates are wrong. All of the variables included in the final model help this process to happen and generate the best estimates for the model that we try to run.

Staff Surveys

Model assumptions:

[38] 42 staff completed baseline interviews, and 29 staff completed baseline and Time 3 interviews, yet 3 interviews were not able to be analyzed, resulting in 26 pairs of interviews analyzed using the PRPAS.

- Each teacher/HV has different amounts of postbaseline measures. The surveys with demographic and baseline information, and at least information from at least one other point in time (T2 or T3 or T4) were included in the final dataset. Observations that did not meet this criterion were excluded from the analysis.
- There are different number of teachers/HVs per site. Thus, we have an unbalanced clustered dataset. LMM can be fitted to this type of data structure with the assumption that any missing data are missing at random.
- The covariates or variables included in the models are:
- Site-level characteristics: type (ECE vs. HV), size, fidelity scores and treatment.
- Teacher/HVs characteristics: dosage, years of experience, gender, age, race, education, an indicator for whether or not they received consultation, and different variables that describe the characteristics of the consultation received (how easy was to schedule the consultation, how valuable was the consultation received, whether or not the consultation received was adequate and the quality of consultation).
- Time-varying variables: time points.
- Dependent variable measures: Maslach Burnout Inventory, Personal Health Questionnaire, Reflective Functioning Questionnaire, Reflective Supervision Rating Scale, Supervisory Working Alliance Inventory, Goal Achievement Scale (for teachers only), and Teacher Opinion Survey (for teachers only)

The staff surveys contained standardized measures of supervisor and staff relationships, reflective functioning, self-efficacy, burnout, and depression. After examining descriptive statistics of our sample, we included the following variables in the models predicting the staff survey outcomes:

- Staff age (2 variables): a categorical variable with 6 categories (under 20 years, 20–29 years, 30–39 years, 40–49 years, 50–59 years and 60 or older); and a categorical variable with 2 categories (under 40 years, 40 or older)
- Staff race/ethnicity (4 variables): a categorical variable with 3 categories (Black, Hispanic, and White); and 3 dummy variables, one for each race category
- Staff educational attainment (2 variables): a categorical variable with 4 categories (Some college/no degree, Associate's degree, Bachelor's degree, Master's degree); and a categorical variable with 2 categories (less than Bachelor's degree, Bachelor's degree or more)
- Program type (center-based or home visiting)
- Program size (small, medium, or large)
- Intervention or Comparison group
- Dosage of intervention

- Time (baseline, Time 2, Time 3, and Time 4)

Classroom Observations

Model Assumptions:

- The classrooms that were observed at baseline and at least one other point in time (T2 or T3 or T4) were included in the final dataset. Observations that did not meet this criterion were excluded from the analysis.
- There are different number of classrooms per site. Thus, we have an unbalanced clustered dataset. LMM can be fitted to this type of data structure with the assumption that any missing data are missing at random.
- The covariates or variables included in the models are:
- Site level characteristics: type (ECE vs. HV), size, fidelity scores and treatment.
- Teacher/HVs characteristics: dosage, years of experience, gender, age, race, education, an indicator for whether or not they received consultation, and different variables that describe the characteristics of the consultation received (how easy was to schedule the consultation, how valuable was the consultation received, whether or not the consultation received was adequate and the quality of consultation).
- Time-varying variables: time points.
- Dependent variable measures: CHILD classroom observation dimension scores

After examining descriptive statistics of our classroom observations sample, we removed variables that were not statistically significant or that were highly correlated with other variables. The following variables were tested in the models predicting the CHILD classroom observation scores:

- Staff age (2 variables): a categorical variable with 6 categories (under 20 years, 20–29 years, 30–39 years, 40–49 years, 50–59 years and 60 or older); and a categorical variable with 2 categories (under 40 years, 40 or older)
- Staff race/ethnicity (4 variables): a categorical variable with 3 categories (Black, Hispanic and White); and 3 dummy variables, one for each race category
- Staff educational attainment (2 variables): a categorical variable with 4 categories (Some college/no degree, Associate degree, Bachelor's degree, Master's degree); and a categorical variable with 2 categories (less than Bachelor's degree, Bachelor's degree or more)
- Program size (small, medium, or large)
- Intervention or Comparison group
- Time

Home Visit Observations

Model Assumptions:

- Home visitors have varying numbers of post-baseline measures.

- There are different numbers of home visitors per site as well as different numbers of families who work with a specific home visitor. Thus, we have an unbalanced clustered dataset. LMM can be fitted to this type of data structure with the assumption that any missing data are missing at random.
- The covariates or variables included in the models are:
- Site-level characteristics: size, fidelity scores and treatment.
- Home visitor characteristics: dosage, years of experience, gender, age, race, education and different staff-level outcomes at four points in time (baseline, Time 2, Time 3 and Time 4).
- Family/Parent-level characteristics: age, race, education, number of children, number of months enrolled in the program, and language.
- Dependent variable: outcomes. We have investigated the effect of the intervention on the following HOVRS scales:
- Home Visitor Responsiveness to Families
- Home Visitor-Family Relationship
- Home Visitor Facilitation of Parent-Child Interaction
- Home Visitor Non-Intrusiveness/Collaboration with Family
- Parent-Child Interaction during Home Visit
- Parent Engagement during Home Visit
- Child Engagement during Home Visit

After examining descriptive statistics of the home visit observations, we removed variables that were not statistically significant or that were highly correlated with other variables. The following variables were tested in the models predicting the HOVRS-A+ scores:

- Staff age (2 variables): A categorical variable with 6 categories (under 20 years old, 20–29 years, 30–39 years, 40–49 years, 50–59 years and 60 or older); and a categorical variable with 2 categories (under 40 years, 40 or older).
- Staff race/ethnicity (4 variables): A categorical variable with 3 categories (Black, Hispanic, and White); and 3 dummy variables, one for each race category
- Staff educational attainment (2 variables): A categorical variable with 4 categories (Some college/no degree, Associate degree, Bachelor's degree, Master's degree); and a categorical variable with 2 categories (less than Bachelor's degree, Bachelor's degree or more)
- Parent age (2 variables): A categorical variable with 6 categories (under 20 years, 20–29 years, 30–39 years, 40–49 years, 50–59 years and 60 or older); and a categorical variable with 2 categories (under 30 years, 30 or older).
- Parent race/ethnicity (4 variables): A categorical variable with 3 categories (Black, Hispanic, and White); and 3 dummy variables, one for each race category

- Parent educational attainment (2 variables): A categorical variable with 4 categories (Less than high school, Some high school, High school/GED, Some college, Bachelor's degree, Master's degree); and a categorical variable with 2 categories (High school or less, More than high school).
- Staff burnout (MBI - Emotional Exhaustion)
- Staff depression (PHQ-2)
- Staff reflective capacity (RFQ - Certainty)
- Staff relationship with supervisors (SWAI - Rapport)
- Staff self-efficacy (TOS and GAS)
- Duration of family enrollment in program, in months
- Intervention or comparison Group
- Time

Child Assessments

Model Assumptions:

- Each student has different amounts of postbaseline measures.
- There are different number of students per classroom as well as different number of teachers within a site. Thus, we have an unbalanced clustered dataset. LMM can be fitted to this type of data structure with the assumption that any missing data are missing at random.
- The covariates or variables included in the models are:
    - Site-level characteristics: size, fidelity scores and treatment.
    - Teacher/HVs characteristics: dosage, years of experience, gender, age, race, and education.
    - Student-level characteristics: gender, age, and race.
    - Time-varying variables: time points.
- Dependent variable: outcomes. We have investigated the effect of the intervention on the following measures:
    - Devereux Early Childhood Assessment (DECA): Total Protective Factors, Initiative, Self-Regulation, Attachment/Relationships, and Behavioral Concerns Scale.
    - Preschool Expulsion Risk Measure (PERM): Total Score, Classroom Disruption Factor, Fear of Accountability Factor, Hopelessness Factor, Teacher Stress Factor.

    - Strengths and Difficulties Questionnaire (SDQ): Total Difficulties Score and Impact Score.

After examining descriptive statistics of the scores on the child assessment scales, we removed variables that were not statistically significant or that were highly correlated with other variables. The following variables were tested in the models for the child assessments:

- Staff age (2 variables): A categorical variable with 6 categories (under 20 years, 20–29 years, 30–39 years, 40–49 years, 50–59 years, and 60 or older); and a categorical variable with 2 categories (under 40 years, 40 or older)
- Staff race/ethnicity (4 variables): A categorical variable with 3 categories (Black, Hispanic, and White); and 3 dummy variables, one for each race category
- Staff educational attainment (2 variables): A categorical variable with 4 categories (Some college/no degree, Associate's degree, Bachelor's degree, Master's degree); and a categorical variable with 2 categories (less than Bachelor's degree, Bachelor's degree or more)
- Child race/ethnicity (4 variables): A categorical variable with 3 categories (Black, Hispanic and White); and 3 dummy variables, one for each race category
- Child age in months
- Child gender
- Staff burnout (MBI - Emotional Exhaustion)
- Staff depression (PHQ-2)
- Staff reflective capacity (RFQ - Certainty)
- Staff relationship with supervisors (SWAI - Rapport)
- Staff self-efficacy (TOS)
- Classroom mental health climate (CHILD Overall Score)
- Intervention or comparison Group
- Time

Appendix B

Table B-1 presents examples of three most frequent categories of issues that consultants in the early childhood center-based and home visiting programs recorded in their logs. Although there were differences between the two types of programs, the most frequent issues were similar. These included (1) issues related to program development and program management; (2) reflective supervision/consultation with staff; and (3) reflective supervision/consultation with supervisors.

Table B-1. Examples of Most Frequent Types of Issues in Consultant Logs

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
Program Development & Management	**Policies & practices:** • Developing a sitewide plan for embedding mental health and social-emotional learning into all aspects of program • Infant and toddler room have not seen increases in development funding source concerned • Reflective supervision with [administrator]. We also discussed continued issues with incomplete files and poor interactions in a specific classroom. • Teachers reflected on policy for sick kids, solutions for dealing with soiled blankets and new clock in policy. **Program expansion**: • Met with director to discuss changes occurring with buying another center. Processed feelings and thoughts about not having	**Policies & practices**: • Discussed program planning for the following fiscal year, including how to be least disruptive to the families being served. Also explored updates on families discussed previously • Explored with individual staff members how they are adjusting to new technology while also becoming more engaged with their families • Goal planning and parent-led interaction • Joined team meeting/group consultation. Themes discussed: worker safety, boundaries with families, thinking around possibly making DCFS hotline call and how to interact with family around this • Did a reflective group with the home visitors and supervisor. This meeting laid out some

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
	the Executive Director here during the transition. Highlighted her strengths and resilience. Checked in with staff. **Staff turnover**: • Concerns about how the work will look and how things will be transitioned with [departure of administrator] from program. • Director turnover. Excelerate [QRIS] questions. • Program strengths and barriers and adjustments anticipated due to administration changes in the program and the building. **Start of school year**: • [Administrator] talked about various changes that might be happening in the program and she updated me on how things went when she did a collaborative joint mission statement exercise with the teachers. • Teacher talked mostly about how registration and administration tasks at the start of the year was stressful. She also indicated that she has a challenging student in her morning class	repeated challenges that staff felt were not being addressed. The focus was in-house. . . what do we need to do. Hopeful. Talked a bit about the diversity issues addressed in the Diversity Informed Tenets. • Thinking about how to use the Piccolo tool with families, how to talk with families about difficult topics such as when home visitors notice negative child behavior **Program expansion**: • Challenges with a growing program - team cohesion, group services. Thinking about roles at work, reflected on supervisors last supervision sessions with home visitors • Met with group to discuss progress in program development, technology, and facilitated discussion on time management and role in their work. **Staff turnover**: • Met with program supervisor to discuss program updates then joined supervisor in reflective supervision with staff. One staff member will be resigning from her position, so there was much conversation about how to manage this and support all families.

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
	who [is having trouble adjusting and] cries a lot.	**Start of school year**: • Attended team meeting with staff as they prepare for the new school year. Discussed technology, recruiting, ongoing monitoring, and reflective supervision needs. Also discussed ongoing management of relationships within the school and how these can impact staff and families. • Discussed programmatic changes coming with the new school year, lessons learned from the first year of the program, and managing stress • Met with full HV staff in two settings- once with full school staff for team building activities and once with just HV group to discuss new families, supervision, program goals, and areas of strength within the program.
Reflective Consultation with Staff	**Administrative issues**: • Coteachers are new to working with issues and having communication issues • Enrollment changes with summer coming, site opening • Met with [staff member] who talked about being frustrated at the process for students who are coming in	**Administrative issues**: • Boundaries with families, staff transitions and expectations, staff engagement and trust • Consultant joined supervisor for her reflective supervision with home visitors. Themes were supporting a less experienced home visitor with her worries in supporting a new family with significant risks, supporting a home

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
	that might need full case studies. • Teacher's health issues that are affecting her ability to do her job, child recently returned to family from foster care has emotional outbursts and boundary issues, mentoring assistant teachers and concerns about buying a new site **Child behavior issues**: • Beginning of new year, separation issues at that time, leaving personal problems at home, past experiences with clingy children. • Children's progress at this stage of the school year, behavior concerns about one child in particular, multiple recent additions to this class • Concerns with one child who prefers to follow the teacher in the classroom and avoids playing with other children. Another child has delays and may have an IFSP or IEP though program was not informed. • Current concerns about one child with some challenging behaviors, her perception of social emotional needs vs academic development. She	visitor who has been struggling with program changes due to growth of the program this year, and supporting a home visitor who is frustrated with a team member. Also general reflection on work with families. • Staff lack of time management • Team cohesion and work styles, developmentally appropriate toys activities, case staffings - mom and baby experiencing separation anxiety, self-care, processing negative responses to work with specific families • Transition families from the program, working with fathers, working with mothers with mental illnesses and addressing leadership boundaries **Child behavior issues**: • Disruptive child with autism, hits home visitor, parents lack of control and attachment with child • Parent concern regarding child self-regulation and mom lack of development knowledge **Family dynamics**: • Arrived to provide consultation to program supervisor and director. However, upon arrival, was informed that a

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
	validated her growing understanding of how her contributions to the children's learning readiness are rooted in social emotional development and successes. • Discussion of ideas to support a child, discussion about engaging a parent • Participated in two reflective supervision sessions with coordinator and TA, and coordinator and para. Discussed concerns with a DD child who appears to be acting out sexually. Explored para's concern around that issue. Discussed children transitioning into a new classroom with TA and the need for children to be on time. Also discussed improvements with separations issues for one child. • Physically and verbally aggressive child - parents are not together and not on the same page regarding how to support child. Teachers struggling with how to work with both parents. • Pre k has some children that tantrum when told "no". The	parent was likely coming in to meet with staff in crisis. Met with the parent and staff members to provide ongoing support. Main issues explored with parent were domestic violence, an order of protection, and shelter. • Transitions with families, relationships with resource representatives • Worker expressed concerns about changes one of her family's is experiencing and the impact that these changes are having on the child. **Parent engagement**: • Consultant joined supervisions with supervisor and home visitors for two separate sessions. Home visitors were able to reflect on their interactions with families. Issues focused around how to partner with families when there are home visitor frustrations or concerns related to scheduling, missed appointments, and parent/child behaviors. • Engaging and supporting parents with mental health issues • How to engage families in having challenging discussions and developing and utilizing goals with families

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
	teachers are new to the room. • Reflective supervision for two classrooms. One room continues to have issues with [child] who does not go to potty. New concern with [child] who teachers suspect of comprehension issues. TA in other room is leaving, making her the fourth to leave in a little over a year. Reflective supervision with teaching team. Teachers reflected on concerns regarding girls in the class bickering and arguing and [child] who tantrums when they don't get what they want. • Reflective supervision with teaching team and supervisor. Reflected on executive functioning and adaptive functioning issues for one child. • Reflective supervision for room. Concerns with separation issues at drop off. One child cries instead of verbalizing need. Other children have language delays. One child who receives ST, OT, DT, is adapting to using the potty in the new classroom.	• Supporting families in developing parent-child goals • Working with families in a preferred language different from the home visitors own primary language **Teacher/Home visitor needs**: • Home visitor ability to self-regulate during visit, difficulty engaging families/parents use of parallel process • Home visitor discussed her work with two families. She discussed her responses/values clashes with family dynamics and feeling frustrated. She shared her struggles with responding to parenting concerns and safety issues in a sensitive and relational manner. She found herself wanting to avoid or shy away from talking about these issues. • Negative feelings around working with families and certain situations, working sensitively within family and household systems. Non-judgmental stance in working with the population - thinking about the culture of poverty.

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
	• Teacher consult following general observation for [child]. We reflected on transitions and issues of inhibitory control • Teacher consult following general observation. Discussed [child] with separation issues and difficulty engaging with others, and second [child] with stranger anxiety. • Teacher consult following individual observation for [child] Teachers concerned about sharing, flexibility and conflict with others. • Teacher consult following observation for [child]. Incident last week where [child] scratched another [child] leaving a mark. • Teachers expressed concerns about helping children calm themselves. • Teaching team reflective supervision. We discussed kids getting bored with toys; laptop, internet and phone extension issues; a mom who is diagnosed with cancer; and children who are being potty trained. • There is a student in a classroom that is hitting and pinching under children. [Their] guardian is rarely in	

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
	communication with the teacher • While [teacher] was on her break, a student started hitting other children. The dean sent him home. Previously, it was decided that children in [this program] would no longer be sent home from school for these behaviors. [Teacher] was uncertain who made the decision to send the child home. She reported that she often feels like she does not know who to take direction from (e.g. principal, superintendent, program director) and often finds that she is told to do many different things by multiple different people. • Family dynamics: Issues with specific family and how their child was doing in classroom • Reflective supervision for [classroom]. Supervision focused on [child] with DD. Parents are separating. We also discussed [child] who has language delays. • Reflective supervision. Child in one has been talking about witnessing DV at home and other issues that may indicate abuse.	

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
	• Reflective for teaching team. We discussed ongoing status of marital issues for [child] with DD. Also discussed how parent separation is affecting a third child. Another child is having attendance issues. **Parent engagement:** • Addressed parent response to event today as 40+ parents attended an interactive parent-child event for year-end fun day. I moved in and out of each group at each activity • Child with speech and trauma concerns and how to engage the mother • Judging families without knowing or understanding their background or story picture. Understanding that all parents want the best for their children are oftentimes are doing the best they can for their children. How to eliminate stress by spreading the staff thin where there is a shortage of staff. • Lack of parent and interest/ involvement in their child's education. Lack of teacher interest/ involvement in participating in education-based activities.	

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
	• Reflective supervision with a lead teacher to discuss recent parent complaints and increase in the use of punishment in the classroom. • Reflective for [teachers in one room]. Reflected on a mom who stays in the classroom for hours due to her anxiety about leaving her child. Also discussed parents who continue to arrive late. Discussed language issues for child who tends to use one-word phrases. Also discussed incomplete files. **Teacher needs (PD, support)**: • Autism and explaining it to typically developing children, communicating difficult issues with parents • Avoiding accidents, all new children, redirecting children, first 45 days, having a class with majority 3-year-olds, issues with referrals to [school district] for speech and social emotional, using ICPS to support social emotional development in the classroom. • [First year teacher's] stress associated with having to move her classroom to	

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
	[another building]. We discussed her feelings and concerns associated with the change. • Identifying and responding to a child's significant learning needs and family limitations in supporting their child's learning. Teacher shared concerns about the depth of needs she sees today and identified she at times has difficulty separating from this concern, i.e. takes it home on weekends and evenings, etc. • Parent communication, aggressive children, difficult transitions, a teacher leaving after being at center for years • Program requirements for further teacher credentialing, numerous recent additions to classroom, teacher stress management for adapting to such changes and the impact on the children's adjustment and learning environment. • Reflective supervision with one classroom. Discussed adapting to changes since TA resignation. Also discussed strategies for	

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
	child following individual observation. • Staff report stress from being short-staffed, director being terminated, and amount of work to be completed at beginning of the year. • Supporting assistant on becoming lead, how to access support • Teachers feeling overwhelmed, the need for better/stronger relationships with the families, children with challenging behavior • Teacher's perspective of her performance to date, concerns and insight about her students and their families, credentialing challenges and her work/life balance. • Teacher's report of her comfort in new school and students' adjustment; external and internal events that impact self-regulation.	
Reflective Consultation with Supervisors	**IECMHC generally**: • Met with the director and assistant director to discuss consultation and how things have been progressing. While the director felt she was seeing progress and liked how things are going,	**IECMHC generally**: • No major issues. Reflected on supervisor's and director's supervisions and interactions with their team. Support for supervisor in her role - feeling some stress with learning this position.

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
	the assistant shared that she still really does not understand what consultation is supposed to look like. • Monthly meeting with two administrators. We discussed creating expectation for monthly reflective supervision with teachers. We also discuss a recent break through with teacher who was able to reflect on conflict with TA. • Pre-supervision reflection with supervisor. Reflected on issues or concerns prior to supervision. • Reflected with supervisor following teachers reflective. We reflected on supervisors' tendency to engage in problem solving for teachers. • Reflection on novel experiences: what kinds of questions to ask to stimulate teaching staff's reflection and problem-solving skills. • Some venting and then discussion about doing workshops with the teachers to support individualizing [instruction] based on child's needs.	**Personnel related stressors**: • Addressed with supervisor concerns about program staff, relationship with director, and recruiting concerns • Building staff capacity for their role, supporting staff confidence via documentation, case consultation process • Feelings of burnout and being overwhelmed by the position duties as supervisor • Program and team changes and transitions

Issue Category	Examples of Issues from Consultant Logs	
	ECE Programs	**HV Programs**
	• Stresses and strains that teachers are feeling and sharing with their coordinator, and how she is responding **Personnel-related stressors**: • Director said that she is challenged with a lead teacher who reportedly is not following simple requests. • Director feeling stressed due to staff conflict and administrative changes • Director stress regarding several teachers calling in sick • Inability to find qualified teachers • Interviewing lead teachers • Supervisors working with challenging employees, teachers engaging with parents whom they do not agree with, parent educator concerned about how to support families in the summer	

Appendix C

Table C-1. Staff Survey Scales - Descriptive Statistics (*N* = 136)

Scale	Group	Baseline Mean *(SD)*	Time 2 Mean *(SD)*	Time 3 Mean *(SD)*	Time 4 Mean *(SD)*
Maslach Burnout Inventory - Emotional Exhaustion (Range 0–54)	Intervention	17.49 (13.22)	16.73 (12.84)	15.05 (12.13)	15.93 (12.62)
	Comparison	15.95 (12.01)	18.90 (12.73)	20.84 (12.91)	17.78 (12.66)
Maslach Burnout Inventory - Depersonalization (Range 0–30)	Intervention	4.07 (5.75)	4.47 (5.41)	3.81 (4.54)	5.07 (6.03)
	Comparison	3.53 (4.74)	6.53 (6.46)	6.02 (5.94)	5.44 (6.29)
Maslach Burnout Inventory - Personal Accomplishment (Range 0–48)	Intervention	35.98 (7.79)	36.28 (8.99)	35.09 (10.64)	35.93 (7.89)
	Comparison	35.78 (8.84)	36.24 (7.28)	35.74 (9.26)	36.02 (9.18)
Personal Health Questionnaire-2 (Range 0–6)	Intervention	0.87 (1.19)	0.71 (1.13)	0.85 (1.21)	0.69 (1.20)
	Comparison	0.95 (1.58)	0.90 (1.23)	1.13 (1.52)	1.09 (1.38)
Reflective Functioning Questionnaire – Certainty (Range 0–3)	Intervention	2.03 (0.77)	1.94 (0.93)	2.11 (0.84)	2.09 (0.86)
	Comparison	1.87 (0.79)	1.86 (0.87)	1.73 (0.90)	1.86 (0.92)
Reflective Functioning Questionnaire – Uncertainty (Range 0–3)	Intervention	0.18 (0.35)	0.26 (0.44)	0.17 (0.41)	0.18 (0.37)
	Comparison	0.20 (0.29)	0.25 (0.33)	0.29 (0.50)	0.29 (0.40)
Reflective Supervision Rating Scale (Range 17–51)	Intervention	42.17 (8.23)	41.22 (9.66)	41.19 (9.60)	41.29 (9.74)
	Comparison	42.62 (8.97)	42.43 (8.49)	39.46 (10.89)	40.64 (10.77)
Supervisory Worker Alliance Inventory- Rapport (Range 1–7)	Intervention	5.69 (1.24)	5.47 (1.52)	5.43 (1.59)	5.47 (1.53)
	Comparison	5.64 (1.28)	5.71 (1.10)	5.38 (1.57)	5.53 (1.33)
Supervisory Worker Alliance Inventory- Client Focus (Range 1–7)	Intervention	5.59 (1.24)	5.39 (1.56)	5.32 (1.61)	5.27 (1.60)
	Comparison	5.53 (1.39)	5.53 (1.23)	5.15 (1.66)	5.33 (1.44)
Goal Achievement Scale – Teachers (Range 0–26)	Intervention	20.98 (3.77)	21.23 (4.18)	20.61 (3.94)	21.67 (3.46)

Scale	Group	Baseline Mean *(SD)*	Time 2 Mean *(SD)*	Time 3 Mean *(SD)*	Time 4 Mean *(SD)*
Goal Achievement Scale – Home Visitors (Range 0–26)	Comparison	21.05 (4.08)	20.58 (4.38)	19.55 (4.86)	21.84 (4.31)
	Intervention	22.07 (2.46)	22.27 (2.41)	22.33 (2.83)	23.14 (2.54)
Teacher Opinion Scale – Teachers (Range 12–60)	Comparison	21.25 (3.40)	23.00 (2.16)	20.58 (5.25)	20.25 (5.91)
	Intervention	46.36 (7.00)	46.59 (7.42)	47.60 (6.71)	46.16 (7.90)
Teacher Opinion Scale – Home Visitors (Range 12–60)	Comparison	47.63 (7.25)	46.37 (6.71)	46.45 (7.30)	48.11 (6.25)
	Intervention	47.80 (4.76)	50.00 (5.34)	46.91 (7.45)	48.29 (5.02)
	Comparison	41.50 (9.00)	40.25 (3.40)	42.50 (11.03)	43.25 (11.47)

Table C-2. Home Visit Rating Scales (HOVRS-A+) Descriptive Statistics (*N* = 7 home visitors, 41 families)

Scale	Group	Baseline Mean *(SD)*	Time 2 Mean *(SD)*	Time 3 Mean *(SD)*	Time 4 Mean *(SD)*
Home Visitor Responsiveness to Family	Intervention	3.60 (0.89)	4.50 (1.00)	3.33 (0.58)	4.50 (0.71)
	Comparison	2.88 (0.84)	2.63 (0.52)	2.50 (0.84)	2.60 (1.34)
Home Visitor Relationship with Family	Intervention	4.40 (1.95)	5.00 (0.82)	5.33 (1.16)	5.50 (2.12)
	Comparison	4.00 (1.60)	3.88 (1.36)	3.83 (1.17)	3.20 (2.17)
Home Visitor Facilitation of Parent-Child Interaction	Intervention	2.40 (1.14)	4.00 (0.82)	2.67 (1.16)	2.50 (0.71)
	Comparison	1.75 (0.89)	2.38 (1.30)	2.50 (1.52)	2.00 (1.41)
Home Visitor Nonintrusiveness/ Collaboration	Intervention	2.20 (0.45)	4.25 (1.26)	3.67 (1.53)	4.00 (1.41)
	Comparison	2.38 (1.30)	2.25 (1.17)	2.50 (1.52)	3.00 (1.58)
Home Visitor Practices Domain	Intervention	3.15 (1.01)	4.44 (0.72)	3.75 (1.00)	4.13 (0.88)
	Comparison	2.75 (0.96)	2.78 (0.99)	2.83 (1.02)	2.70 (1.54)
Parent-Child Interaction	Intervention	3.80 (1.10)	5.50 (0.58)	5.67 (0.58)	5.50 (2.12)
	Comparison	3.88 (1.13)	4.00 (1.69)	4.17 (2.04)	5.80 (1.30)
Parent Engagement	Intervention	4.60 (1.82)	5.25 (0.96)	5.67 (1.53)	5.50 (2.12)
	Comparison	4.25 (0.71)	4.50 (1.41)	4.00 (1.10)	5.20 (1.48)
Child Engagement	Intervention	4.60 (2.07)	5.00 (2.31)	6.00 (1.00)	4.50 (2.12)
	Comparison	4.50 (1.69)	4.25 (1.83)	4.50 (2.51)	5.80 (1.64)
Family Engagement	Intervention	4.33 (1.55)	5.25 (1.26)	5.78 (1.02)	5.17 (2.12)
	Comparison	4.21 (1.05)	4.25 (1.39)	4.22 (1.61)	5.60 (1.38)
IECMHC scale	Intervention	3.29 (0.93)	4.88 (0.24)	4.05 (1.15)	4.42 (1.03)
	Comparison	2.82 (0.98)	3.01 (1.13)	3.22 (1.43)	3.51 (1.80)

Table C-3. Child Assessment Scales – Descriptive Statistics (*N* = 136)

Scale	Group	Baseline Mean *(SD)* (*N* = 136)	Time 2 Mean *(SD)* (*N* = 58)	Time 3 Mean *(SD)* (*N* = 54)
DECA[a] – Initiative	Intervention	20.57 (5.97)	23.39 (6.30)	23.87 (5.77)
	Comparison	20.74 (4.89)	22.31 (6.66)	24.08 (6.61)
DECA – Self-Regulation	Intervention	21.70 (7.16)	21.87 (5.87)	22.44 (5.56)
	Comparison	22.51 (6.04)	21.54 (7.71)	22.38 (6.51)
DECA – Attachment/Relationships	Intervention	25.07 (4.97)	25.21 (4.67)	24.67 (5.00)
	Comparison	26.71 (4.58)	26.69 (5.07)	25.75 (4.69)
DECA – Total Protective Factors[b]	Intervention	141.79 (22.01)	144.08 (24.45)	146.14 (23.79)
	Comparison	145.92 (20.47)	147.15 (28.61)	144.82 (25.48)
DECA – Behavioral Concerns	Intervention	12.69 (7.35)	10.68 (5.33)	10.24 (5.19)
	Comparison	12.48 (7.88)	13.83 (9.62)	12.31 (8.90)
Preschool Risk Expulsion Measure – Total Score[c]	Intervention	7.14 (3.65)	5.79 (2.78)	6.01 (2.84)
	Comparison	7.15 (3.69)	6.82 (3.81)	6.79 (4.10)
Preschool Risk Expulsion Measure – Classroom Disruption	Intervention	2.04 (1.20)	1.56 (0.94)	1.63 (0.95)
	Comparison	1.91 (1.14)	1.85 (1.14)	1.90 (1.29)
Preschool Risk Expulsion Measure – Fear of Accountability	Intervention	1.77 (1.06)	1.38 (0.56)	1.53 (0.79)
	Comparison	1.75 (1.07)	1.84 (1.27)	1.79 (1.16)
Preschool Risk Expulsion Measure – Hopelessness	Intervention	1.62 (0.76)	1.41 (0.69)	1.44 (0.67)
	Comparison	1.65 (0.83)	1.59 (0.80)	1.67 (0.92)
Preschool Risk Expulsion Measure – Teacher Stress	Intervention	1.68 (0.99)	1.43 (0.85)	1.35 (0.73)
	Comparison	1.66 (0.97)	1.54 (0.90)	1.44 (0.90)
Strengths and Difficulties Questionnaire – Any Difficulties, %(*n*)[d]	Intervention	55% (*n* = 48)	44% (*n* = 20)	48% (*n* = 19)
	Comparison	49% (*n* = 24)	62% (*n* = 8)	54% (*n* = 7)
Strengths and Difficulties Questionnaire – Impact Score[e]	Intervention	2.12 (1.96)	1.45 (1.39)	1.11 (1.63)
	Comparison	2.09 (1.64)	2.38 (2.39)	2.29 (2.06)

[a] Devereux Early Childhood Assessment for Preschoolers, Second Edition. Possible range of raw scores for Initiative, Self-Regulation, Attachment/Relationship, and Behavioral Concerns scales is 0–36.

[b] Total Protective Factors is calculated by summing the t-scores from the Initiative, Self-Regulation, and Attachment/Relationship scales. Raw scores for these scales are presented in the table.

[c] Preschool Risk Expulsion Measure Total score is the sum of the subscale scores; possible range for total score is 4–20. Subscale scores (Classroom Disruption, Fear of Accountability, Hopelessness, Teacher Stress) are the means of the item responses, ranging from 1 (strongly disagree) to 5 (strongly agree).

[d] SDQ item: "Overall, do you think that your child has difficulties in one or more of the following areas: emotions, concentration, behavior or being able to get on with other people?" Percent of children for whom the teacher responded "yes" are presented in the table.

[e] SDQ Impact score items were only administered to teachers who responded "yes" to the SDQ item asking if the child has any difficulties (n = 71 at baseline, n = 28 at Time 2, n = 26 at Time 3). SDQ Impact Score is calculated by summing the responses for the items assessing the impact of the difficulties on the child, peer relationships, and learning. Possible range for Impact score: 0–6. Score of 0=Near average, 1=Slightly raised, 2=High, and 3+=Very high.

Table C-4. Parent Survey Scales – Descriptive Statistics (*N* = 51)

	Baseline		Time 2		Time 3		Time 4				
Scale	Intervention (*N* = 14) Mean (*SD*)	Comparison (*N* = 8) Mean (*SD*)	Intervention (*N* = 4) Mean (*SD*)	Comparison (*N* = 8) Mean (*SD*)	Intervention (*N* = 4) Mean (*SD*)	Comparison (*N* = 6) Mean (*SD*)	Intervention (*N* = 2) Mean (*SD*)	Comparison (*N* = 5) Mean (*SD*)	F[a]	p	η_p^2
HFPI - Problem Solving	24.00 (3.01)	24.38 (3.16)	25.00 (1.15)	24.13 (1.36)	25.75 (2.22)	26.00 (4.15)	25.00 (7.07)	21.80 (5.89)	1.82	.088^	.41
HFPI- Parent/Child Interaction	44.71 (4.25)	44.38 (4.14)	47.00 (2.16)	44.75 (3.65)	47.25 (3.77)	46.17 (1.60)	50.00 (0.00)	47.40 (4.22)	1.14	.369	.31
HFPI- Home Environment	34.64 (5.17)	35.13 (3.14)	38.25 (2.87)	32.25 (5.90)	35.00 (5.03)	34.83 (2.99)	37.00 (1.41)	36.40 (4.77)	1.28	.277	.33
HFPI - Role Satisfaction	20.93 (2.46)	20.50 (2.00)	21.25 (0.96)	20.00 (1.93)	21.25 (5.12)	20.50 (2.17)	24.00 (0.00)	19.60 (1.82)	0.44	.933	.15
HFPI - Parenting Efficacy	25.29 (4.39)	25.75 (3.58)	25.50 (3.00)	26.75 (3.54)	24.50 (3.87)	26.83 (3.60)	29.50 (0.71)	28.60 (3.13)	1.00	.472	.28
PHQ - Depression	1.50 (2.14)	0.00 (0.00)	1.00 (1.41)	0.63 (1.19)	1.00 (1.15)	1.67 (2.42)	1.50 (2.12)	0.80 (1.79)	1.15	.361	.31
ICPW - Perception of Infant Crying	(*n* = 4)[b] 2.25 (0.50)	(*n* = 3) 2.00 (3.46)	(*n* = 2) 0.00 (0.00)	(*n* = 3) 0.67 (1.15)	(*n* = 3) 2.00 (2.83)	(*n* = 1) 1.00	(*n* = 1) 0.00	(*n* = 2) 0.00 (0.00)	--	--	--

[a] The F-statistics presented are results from between-subjects interaction effects of time by group, with program nested within group.

[b] The scale used to assess parental perception of infant crying (ICPW) was only administered to parents with infants 12 months of age or younger, hence the smaller sample size for this measure. The ICPW was not analyzed by group over time due to the very small sample.

Table C-5. Summary of Results from Analyses of Standardized Measures

Measure – subscale	Type of analytic model	Expected direction of effect of intervention	Significance level of intervention effect	If significant, did it match the expected direction?
Staff survey				
Reflective Functioning Questionnaire - Certainty	Linear mixed modeling (LMM)	+	^	Yes
Reflective Functioning Questionnaire - Uncertainty	LMM	+	*	Yes
Maslach Burnout Inventory – Emotional Exhaustion	LMM	-	N.S.	--
Maslach Burnout Inventory – Depersonalization	LMM	-	N.S.	--
Maslach Burnout Inventory – Personal Accomplishment	LMM	+	N.S.	--
Personal Health Questionnaire	LMM	-	N.S.	--
Goal Achievement Scale	LMM	+	Model did not run	--
Teacher Opinion Scale	LMM	+	N.S.	--
Reflective Supervision Rating Scale	LMM	+	N.S.	--
Supervisory Worker Alliance Inventory – Rapport	LMM	+	N.S.	--
Supervisory Worker Alliance Inventory – Client focus	LMM	+	N.S.	--
Provider Reflective Practice Assessment Scales				
Self-Knowledge	Repeated measures analysis of variance (ANOVA)	+	N.S.	--
Self-Regulation	Repeated measures ANOVA	+	N.S.	--
Multiple Perspectives	Repeated measures ANOVA	+	*	Yes
Collaboration	Repeated measures ANOVA	+	N.S.	--
Process	Repeated measures ANOVA	+	N.S.	--
Authentic Attitude	Repeated measures ANOVA	+	N.S.	--

Measure – subscale	Type of analytic model	Expected direction of effect of intervention	Significance level of intervention effect	If significant, did it match the expected direction?
Classroom observations – Climate of Healthy Interactions for Learning & Development				
Transitions	LMM	+	N.S.	--
Directions & Rules	LMM	+	*	Yes
Social & Emotional Learning	LMM	+	N.S.	--
Staff Awareness	LMM	+	N.S.	--
Staff Affect	LMM	+	N.S.	--
Staff Cooperation	LMM	+	Model did not run	--
Staff-Child Interactions	LMM	+	N.S.	--
Individualized & Developmentally Appropriate Pedagogy	LMM	+	^	Yes
Child Behaviors	LMM	+	N.S.	--
Equity	LMM	+	*	Yes
Warmth	LMM	+	N.S.	--
Home Visit Observation Rating Scales (HOVRS-A+)				
Responsiveness to Family	LMM	+	*	Yes
Relationship with Family	LMM	+	N.S.	
Facilitation of Parent-Child Interaction	LMM	+	^	Yes
Nonintrusiveness/Collaboration	LMM	+	N.S.	--
Home Visitor Practices Domain	LMM	+	N.S.	--
Parent-Child Interaction	LMM	+	N.S.	--
Parent Engagement	LMM	+	N.S.	--
Child Engagement	LMM	+	N.S.	--
Family Engagement Domain	LMM	+	N.S.	--
IECMHC scale	Two-way ANOVA	+	*	Yes
Child assessments				
Strengths & Difficulties Questionnaire – Impact Score	LMM	-	^	Yes

Measure – subscale	Type of analytic model	Expected direction of effect of intervention	Significance level of intervention effect	If significant, did it match the expected direction?
Preschool Expulsion Risk Measure – Total Score	LMM	-	N.S.	--
Preschool Expulsion Risk Measure – Classroom Disruption	LMM	-	N.S.	--
Preschool Expulsion Risk Measure – Fear of Accountability	LMM	-	N.S.	--
Preschool Expulsion Risk Measure – Hopelessness	LMM	-	N.S.	--
Preschool Expulsion Risk Measure – Teacher Stress	LMM	-	N.S.	--
DECA – Total Protective Factors	LMM	+	N.S.	--
DECA – Initiative	LMM	+	N.S.	--
DECA – Self-Regulation	LMM	+	N.S.	--
DECA – Attachment/Relationships	LMM	+	N.S.	--
DECA – Behavioral Concerns	LMM	-	N.S.	--
Parent survey				
Healthy Families Parenting Inventory – Problem Solving	Analysis of variance (ANOVA)	+	N.S.	--
Healthy Families Parenting Inventory – Parent/Child Interaction	ANOVA	+	N.S.	--
Healthy Families Parenting Inventory – Home Environment	ANOVA	+	N.S.	--
Healthy Families Parenting Inventory – Role Satisfaction	ANOVA	+	*	Yes
Healthy Families Parenting Inventory – Parenting Efficacy	ANOVA	+	N.S.	--
Personal Health Questionnaire	ANOVA	-	N.S.	--

^$p < .10$, *$p < .05$, **$p < .01$, ***$p < .001$

Note. Our usual p level for statistical significance is < .05. However, we also report results of $p < .10$ to note a trend in the data.

Brief descriptions of measures:

The Reflective Functioning Questionnaire (RFQ; Fonagy et al., 2016) consists of 8 different items which are re-scored to be used in two different subscales; each consisting of 6 items. The initial scale ranges from 1 ("Strongly disagree") to 7 ("Strongly agree"). Each of the 8 items is re-scored into a scale ranging from 0 to 3. The 6 items in each subscale are averaged together so that the Certainty subscale would have a possible score range of 0–3 and the Uncertainty subscale would have a possible score ranging from 0–2.33.

The Maslach Burnout Inventory (MBI)-Human Services (Maslach et al., 1996), contains 22 items measuring three facets of burnout in the following subscales: Emotional Exhaustion, Depersonalization, and Personal Accomplishment. The response scale, which is labeled at each point, ranges from 0 ("Never") to 6 ("Every day"). Subscale scores are sums of the item scores, resulting in possible scores ranging from 0 to 54 for Emotional Exhaustion (9 items), 0 to 30 for Depersonalization (5 items), and 0 to 48 for Personal Accomplishment (8 items).

The Personal Health Questionnaire (PHQ; Kroenke et al., 2003) is a measure of depression that consists of two items. The response scale, which is labeled at each point, ranges from 0 ("Not at all") to 3 ("Nearly every day"). The two items are summed, resulting in possible scores ranging from 0 to 6.

The Goal Achievement Scale (GAS; Alkon et al., 2003) consists of 13 items (a 14th item was not included in our analysis because it could not be asked at baseline). The response scale, which is labeled at each point, ranged from 0 ("Not at all") to 2 ("Very much"). The 13 items are summed, resulting in possible scores ranging from 0 to 28. With the authors' permission, we adapted the GAS to administer to home visitors.

The Teacher Opinion Scale (TOS; Geller & Lynch, 1999) consists of 12 items. The response scale, which is labeled at each point, ranges from 1 ("Strongly disagree") to 5 ("Strongly agree"). The 12 items are summed together, resulting in possible scores ranging from 12 to 60. With the authors' permission, we adapted the TOS to administer to home visitors. The TOS was not administered to supervisors.

The Reflective Supervision Rating Scale (RSRS; Ash, 2010) consists of 17 items. The response scale, which is labeled at each point, ranged from 1 ("Rarely") to 3 ("Almost always"). The 17 items are summed, resulting in possible scores ranging from 17 to 51. The RSRS was not administered to supervisors.

The Supervisory Worker Alliance Inventory (SWAI; Efstation et al., 1990) consists of 19 items for frontline staff and supervisors. These 19 items are used in two subscales: the Rapport subscale consisting of 12 items and the Client Focus subscale consisting of 7 items. The supervisor version asks respondents to complete 7 additional items for the Identification subscale. The response scale ranged from 1 ("Almost never") to 7 ("Almost always"). Each subscale is an average of the items in it, thus possible scores range from 1 to 7 for each subscale.

The Provider Reflective Process Assessment Scales (PRPAS; Heller, 2017) consists of 14 items in six scales that correspond to different dimensions of reflective capacity: Self-Knowledge (2-item scale); Self-Regulation (3-item scale); Multiple Perspectives (1-item scale); Collaboration (3-item scale); Process (2-item scale), and Authentic Attitude (3-item scale). The response scale ranges from 0 ("Reverse") to 4 ("High"). The items in each scale are summed, resulting in possible scores ranging from 0 to 4 or 0 to 12 depending on the number of items in each scale.

The Climate of Healthy Interactions for Learning and Development (CHILD; Gilliam & Reyes, 2017) tool is an observational assessment of the social and emotional (mental health) climate of early care and education settings. CHILD items are scored on a 5-point scale ranging from -2 to +2. Negative scores indicate a climate that undermines a child's mental health; positive scores indicate one that promotes mental health; and 0 sets the baseline expectation of "doing no harm." The CHILD consists of 28 behavioral items across nine dimensions: Transitions, Directions & Rules, Social & Emotional Learning, Staff Awareness, Staff Affect, Staff Cooperation, Staff-Child Interactions, Individualized & Developmentally Appropriate Pedagogy, and Child Behaviors. The measure developers also created two auxiliary scales comprised of items from other scales: Equity and Warmth.

The Home Visit Rating Scales - Adapted & Extended to Excellence (HOVRS-A+; Roggman et al., 2010) is a widely used observation tool to assess home visitors' strategies and relationships during home visits. It consists of the following seven scales rated from 1 ("Inadequate") to 7 ("Excellent"), with higher ratings reflecting more responsive behaviors on the part of the home visitor: Home Visitor Responsiveness to Family, Home Visitor Relationship with Family, Home Visitor Facilitation of Parent–Child Interaction, Home Visitor Non-Intrusiveness & Collaboration, Parent–Child Interaction during Home Visit, Parent Engagement during Home Visit, Child Engagement during Home Visit.

We used a brief, modified version of the Strengths & Difficulties Questionnaire (SDQ; Goodman et al., 2003) containing six items (Perry, 2013). Only respondents who indicated that a child has difficulties with emotions, concentration, behavior, or getting along with other people were administered the remaining five items to assess the severity of the problems, with a response scale of 0 ("Less than a month"/"Not at all") to 2 ("Over a year"/"A great deal"). Three items asking if these difficulties upset the child, if these difficulties interfere with the child's everyday life

in peer relationships, and if these difficulties interfere with the child's learning are summed, resulting in an Impact Score, which can range from 0 to 6.

The Preschool Expulsion Risk Measure (PERM; Gilliam & Reyes, 2018) is a 12-item tool that uses a 5-point scale ranging from 1 ("Strongly disagree") to 5 ("Strongly agree"). It measures teacher perception of disruptive child behavior in four subscales: Classroom Disruption (the degree to which a child's behaviors create disruptions in the classroom); Fear of Accountability (the degree to which children's behaviors may pose a risk of injury for which the teacher might be accountable); Hopelessness (the degree to which the teacher may feel hopeless that anything can be done to improve behaviors in the classroom); and Teacher Stress (the degree to which children's behaviors are associated with increased teacher stress). The PERM also provides a Total Score.

The Devereux Early Childhood Assessment (DECA-P2; LeBuffe & Naglieri, 2012) version for ages 36 to 60 months includes three scales assessing protective factors: Initiative, Self-Regulation, and Attachment/Relationships, resulting in a Total Protective Factors scale. It also includes a Behavioral Concerns scale, which assesses challenging behaviors such as aggression, withdrawal, inattention, and exhibiting extreme emotions. This measure includes 38 items which are scored on a 4-point scale ranging from 0 ("No, none") to 3 ("Yes, severe").

Five subscales from the Healthy Families Parenting Inventory (HFPI; LeCroy, Krysik, & Milligan, 2007) were included in the parent survey: Parent/Child Interaction, Home Environment, Role Satisfaction, Parenting Efficacy, and Problem Solving. Items contain response options from 1 ("Rarely or never") to 5 ("Always or most of the time"). The HFPI was specifically developed to measure families in home visiting programs, and each subscale can be administered on its own (Krysik & LeCroy, 2012).

Table C-6. Pearson Correlations of Change in Reflective Functioning (RFQ) and Final Burnout Levels (*N* = 86)

	1	2	3	4	5
RFQ Certainty – change from T1-T4	-				
RFQ Uncertainty – change from T1-T4	-.53**	-			
MBI Emotional Exhaustion T4	-.01	.13	-		
MBI Depersonalization T4	-.06	.07	.68**	-	
MBI Personal Accomplishment T4	.10	.07	.03	-.02	-

**$p < .01$

Table C-7. Pearson Correlations of Change in Reflective Practice (PRPAS) and Final Burnout Levels (*N* = 23)

	1	2	3	4	5	6	7	8	9
PRPAS Self-Knowledge – change from T1 to T3	-								
PRPAS Self-Regulation – change from T1 to T3	.59**	-							
PRPAS Multiple Perspectives – change from T1 to T3	.46*	.22	-						
PRPAS Collaboration – change from T1 to T3	.36	.46*	.30	-					
PRPAS Process – change from T1 to T3	.55**	.39*	.23	.73***	-				
PRPAS Authentic Attitude – change from T1 to T3	.51**	.46*	.33	.61***	.63**	-			
MBI Emotional Exhaustion T3	-.34	-.19	-.33	-.48*	-.49*	-.29	-		
MBI Depersonalization T3	-.03	.29	-.24	-.05	.08	-.15	.60**	-	
MBI Personal Accomplishment T3	.27	.02	.13	.10	-.12	.15	-.11	-.18	-

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

Table C-8. Follow-up Parent Sample Survey Change Score Analysis (*N* = 21)

	Group	N	Mean	*SD*	T
HFPI Problem Solving – Change score	Comparison	12	1.92	3.09	0.59
	Intervention	9	1.00	4.06	
HFPI Role Satisfaction – Change score	Comparison	12	0.17	3.13	1.14
	Intervention	9	-1.44	3.32	
HFPI Parent Child Interaction – Change score	Comparison	12	-0.25	5.88	0.14
	Intervention	9	-0.56	3.43	
HFPI Home Environment – Change score	Comparison	12	-0.17	4.63	0.67
	Intervention	9	-1.56	4.85	
HFPI Parenting Efficacy – Change score	Comparison	12	0.00	4.45	0.39
	Intervention	9	-0.78	4.58	
PHQ Score – Change score	Comparison	12	0.00	0.85	1.36
	Intervention	9	-0.56	1.01	

Note. No significant group differences were found between change scores.