Indiana Maternal Infant Early Childhood Home Visiting (IN MIECHV): FY2018 Evaluation Report

Evaluation of Healthy Families Indiana (HFI) Mental Health Consultation Enhancement

October 2020

DIEHL CONSULTING GROUP
evaluation | analytics | solutions

This Indiana Maternal Infant Early Childhood Home Visiting (IN MIECHV): FY2018 Evaluation Report was supported by the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services (HHS) as part of an award totaling $10,911,705.00 with zero percentage financed with nongovernmental sources. The contents are those of the author(s) and do not necessarily represent the official views of, nor an endorsement, by HRSA, HHS or the U.S. Government.

Table of Contents

Introduction and Project Background 4
VI. Evaluation Summary 4
A. & B. Evaluation Questions, Study Design, and Target Population 4
C. Major Findings 5
D. Limitations 7
E. Implications of Evaluation Findings 8
F. Lessons Learned 9
VII. Evaluation Design 9
A. Entities/Organizations Responsible for Collection and Reporting Evaluation Data 9
B. Evaluation Rationale 10
C. Description of Adaptation/Enhancement 11
D. Use of Prior Evaluation Findings 11
E. Theory of Change and Evaluation Framework 13
F. Outcomes 13
G. Target Populations 14
H. Evaluation Questions 14
I. Evaluation Design 15
J. Rationale for Design 16
K. Evaluation Timeline 17
L. Instruments 18
M. Design Specific Components 21
VIII. Data Collection Methods and Schedule 21
A. Data Collection 21
B. Design Specific Components 22
IX. Sample Size and Sampling Plan 22
A. Sampling Plan 22
B. Statistical Power 22
C. Design Specific Components 23
X. Analytic Methods 24
A. Analytic Methods 24
B. Evaluation Cost 26
C. Design Specific Components 26
XI. Evaluation Results 27

A. Results 27
B. Sample and Comparison Group 54
C. Comparison Group 59
D. Discussion and Interpretation 59
E. Limitations 65
XII. Evaluation Successes and Challenges, Conclusions, and Dissemination of Evaluation Findings 67
A. Strategies That Facilitated Evaluation Implementation 67
B. Evaluation Successes 67
C. Challenges 68
D. Adherence to Proposed Plan 68
E. Key Findings 68
F. Implications of Findings 69
G. Recommendations 70
H. Detailed Plan for Dissemination 71
References 73

Introduction and Project Background

The Maternal, Infant, and Early Childhood Home Visiting (MIECHV) program supports at-risk families by providing resources to ensure that children are physically, socially, and emotionally healthy and ready to learn. The purpose of Indiana MIECHV is to support the delivery of coordinated and comprehensive high-quality voluntary early childhood home visiting services to eligible families. The overall vision of Indiana MIECHV is to improve health and developmental outcomes for children and families who are at risk. Mental Health Consultation (MHC) is an enhancement to the Healthy Families Indiana (HFI) program providing licensed mental health clinicians tasked with supporting home visitors' and families' mental health by monitoring family records, supporting home visitors with strategies, identifying overall trends, and providing relevant training.

VI. Evaluation Summary

A. & B. Evaluation Questions, Study Design, and Target Population

The FY2018 evaluation was designed to 1) identify the supporting factors and barriers associated with implementing MHC with fidelity at the site level, 2) examine the effects of perceived MHC fidelity on staff perceptions of quality/relevance/usefulness of resources, self-efficacy, secondary trauma, compassion satisfaction, and burnout, and 3) explore the effects of MHC services on parenting and family functioning/support using a dose-response framework that considers increased fidelity to the treatment model as an increased dose.

Table 1. Summary of research questions, study design, and target population.

A1. Research Questions	A2. Study Design	B. Target Population
RQ1. What are the supporting factors and barriers associated with implementing Mental Health Consultation with fidelity to the Mental Health Consultation model?	To address research question 1, an implementation-focused evaluation (CFIR, 2019; Stetler et al., 2006) utilizing a qualitative design and semi-structured interviews was employed.	Interview responses were drawn from the full population of 1) HFI Program Managers and 2) Mental Health Consultants.
RQ2. To what extent are home visitor perceptions of Mental Health Consultation fidelity associated with ratings of perceived quality/relevance/usefulness of resources, self-efficacy, secondary trauma, compassion satisfaction, and burnout?	An exploratory, correlational design was employed to examine the relationship between home visitors' ratings of MHC fidelity and ratings of perceived quality/relevance/usefulness of resources, self-efficacy, secondary trauma, compassion satisfaction, and burnout.	Responses were drawn from the full population of home visitors who received MHC at HFI sites serving MIECHV-funded families (using the MIECHV MHC model).
RQ3. What is the effect of Mental Health Consultation on parenting and family functioning/support outcomes as measured by the depression, personal care, and mobilizing resources subscales of the Healthy Families Parenting Inventory (HFPI) and the emotional and verbal responsivity of primary caregiver, organization of physical and temporal environment, and parent involvement with child subscales from the Home Observation for Measurement of the Environment Inventory (HOME)? (i.e., 1) Among the families with Mental Health Consultation, do those with higher fidelity to the MHC model have better outcomes; and 2) Do families receiving MHC have better outcomes than families not receiving the enhancement?)	The evaluation team 1) determined the objective and subjective measures of fidelity that were related to family outcomes, 2) created a summary measure of fidelity for each family drawn from measures of fidelity shown to be predictive of outcomes, and 3) examined whether the fidelity score predicts family outcomes within the MHC treatment group. In a second step, an exploratory quasi-experimental matched comparison groups design was employed. MIECHV-funded families served by home visitors receiving the MHC enhancement were divided into low and high fidelity treatment groups, and each was compared to a separate matched group of non-MIECHV-funded families served by home visitors not receiving the MHC enhancement.	The matched groups were drawn from the full population of families participating in the HFI program from January 1, 2016 to December 30, 2019 (*N* = 19,302 families) for whom outcome data were available (*N* = 7,943 families). This consisted of 1,692 MIECHV-funded families receiving the MHC enhancement and 6,251 non-MIECHV-funded families. Families who enrolled in HFI prior to implementation of the Localized MHC model were excluded from the sample. Non-MIECHV-funded families (not receiving MHC) in sites serving MIECHV-funded families (receiving MHC) were excluded from the comparison groups. Families in Decatur County and the HFI site in Marion County that does not serve MIECHV-funded families were excluded because they implemented a different model of MHC.

C. Major Findings

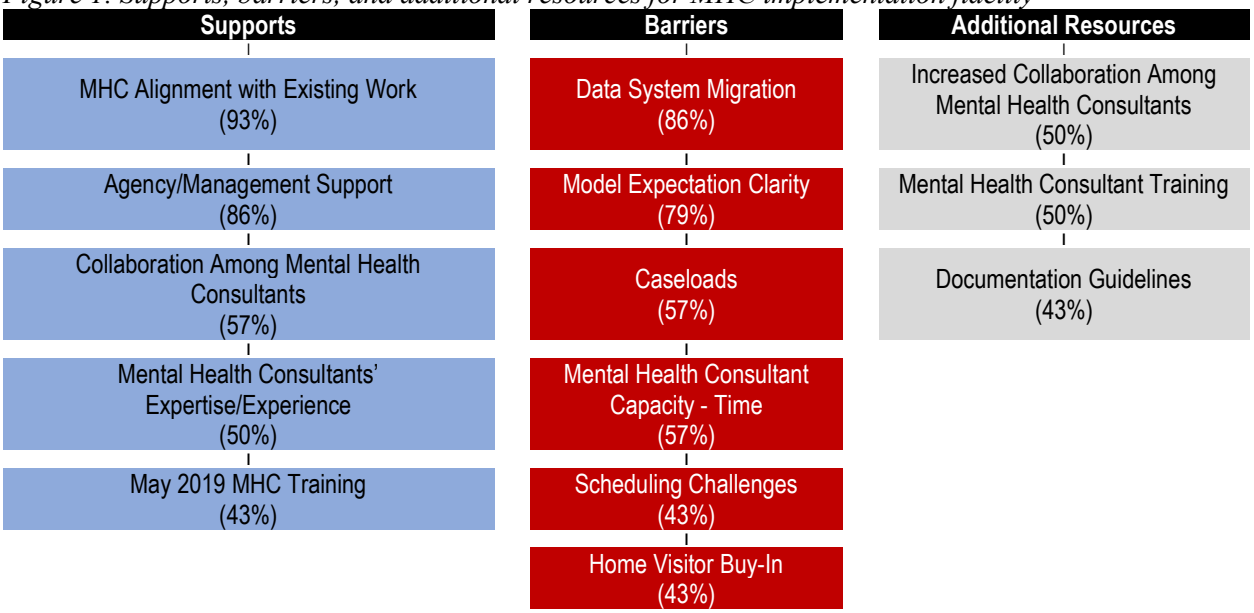
Research Question 1

During August and September 2019, Diehl Consulting Group (DCG) completed semi-structured interviews with mental health consultants and program managers. Fourteen staff members (7 mental health consultants and 7 program managers) participated in the study.

Implementation Discrepancies. Overall, mental health consultants and program managers reported that the majority of model expectations were met; however, responses suggested that most sites experienced challenges meeting documentation expectations and that some sites struggled to complete monthly reviews for all MIECHV-funded families. Furthermore, the majority of sites noted that monthly reviews were not completed and/or documented in adherence to model expectations prior to May 2019.

Supports, Barriers, and Additional Resources. Key themes (i.e., those described by at least 40% of participants) were identified using the framework method. Supports, barriers, and additional resources identified during the analysis are included below.

Figure 1. Supports, barriers, and additional resources for MHC implementation fidelity

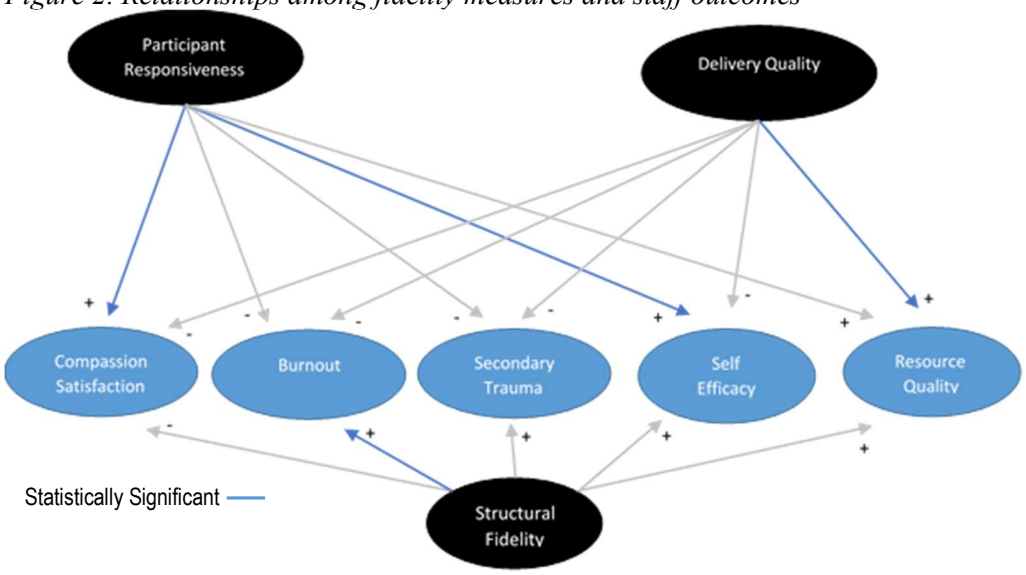


Research Question 2

To examine the relationships between home visitors' ratings of MHC fidelity and ratings of perceived quality/relevance/usefulness of resources, self-efficacy, secondary trauma, compassion satisfaction, and burnout, a partial least squares-path analysis (PLS-PA) was utilized. A fully linked model was created that explored the unique relationships between each fidelity measure and each staff outcome. The model suggested that the individual aspects of fidelity make different, at least partially unique contributions to the home visitor outcomes. There was a large, significant relationship between delivery quality and perceived quality of the resources. Increased participant responsiveness (i.e., home visitors' confidence completing key aspects of MHC) was associated with greater self-efficacy (a medium-sized effect) and greater compassion satisfaction (a small effect) among the home visitors. An increase in adherence to the structural

aspects of fidelity (e.g., model adherence and model exposure) was associated with an increase in burnout (a small effect). This latter relationship suggests that home visitors may find participating in MHC to be demanding and burdensome when it is implemented with fidelity to the model. However, because this was a small effect, these results suggest a need to further explore the obstacles home visitors may encounter.

Figure 2. Relationships among fidelity measures and staff outcomes



Research Question 3

RQ3a. The results suggested that fidelity to the MHC treatment model predicted family outcomes within the MHC treatment group for the HOME outcomes, but not for the HFPI outcomes. For the HOME subscales (Responsivity, Organization, and Involvement), there were effects of total fidelity score on the slopes of outcome measure, which indicated greater improvement in the outcomes as fidelity increased. Effects were very small, with fidelity accounting for 1% of the variance in the change in the HOME outcomes over time. For the HFPI subscales (Mobilizing Resources, Personal Care, Depression), no effects of fidelity on the change in these subscales were observed. The main analyses were supplemented by exploratory analyses that compared the mean change over time in each of six subscales for two groups composed of MHC families with the highest and lowest fidelity scores. These analyses confirmed the findings from the primary analyses. In summary, RQ3a results suggest an effect of fidelity to the MHC model on HOME outcomes, albeit a small one.

RQ3b. For High Fidelity MHC families, no differences in improvement were observed when compared to their matched comparison group. For Low Fidelity MHC families, significant differences were observed, with Low Fidelity MHC families showing less improvement compared to their matched comparison group. On the HOME subscales (Responsivity, Organization, and Involvement), all groups tended to improve over time except the Low Fidelity MHC treatment group, which either did not improve over time or demonstrated less improvement than other groups. On the HFPI subscales (Mobilizing Resources, Personal Care, Depression), the data suggested no changes over time for all groups except the Low Fidelity treatment group, which actually worsened over time on these subscales. Two possible interpretations have emerged from these results: 1) Because the majority of the total fidelity score was determined by factors occurring at the site level rather than at the family level, it is

possible that poor performing sites both have difficulty adhering to the MHC model with fidelity and execute their invention more poorly. 2) Within the dose-response framework used for the evaluation, increased fidelity to the treatment model is interpreted as an increased treatment dose. Therefore, an alternative possibility is that because MIECHV sites are specifically selected due to their location in high-risk counties, Low Fidelity MHC families may represent higher-risk families that essentially did not receive a treatment (based on the theorized dose-response relationship that defines lower fidelity as a lower treatment dose). This may suggest that families from high-risk communities not receiving a treatment, or receiving a low-quality treatment, may perform more poorly in general than those from lower risk communities. In this case, the fact that the High Fidelity MHC treatment group performed the same as its matched-comparison group may indicate that it is actually performing much better than it otherwise would have without the treatment because it is composed of families from high-risk counties.

D. Limitations

RQ1. Evaluation data were drawn from self-report interviews, which can create a number of limitations including, but not limited to, social desirability bias, attribution issues, and memory errors. Moreover, responses reflected the perceptions of the interviewee and may not be independently verifiable. Finally, responses reflect each participant's understanding of specific MHC model expectations and may have varied across respondents and/or have been inconsistent with the MHC contractual expectations. As applicable, efforts have been made throughout this report to triangulate interview responses with other data sources for the most accurate interpretations. **RQ2.** Due to the small sample size, the analysis was "rank deficient" (i.e., there were more measurement items (88) than observations (N = 66)). As a result, the overall analysis is not as reliable as it would be were it not rank deficient; therefore, some caution in interpretation is necessary. In particular, one should use caution when interpreting any small effects (both their size and direction) because small effects in a small sample may not be present or even reverse direction in a larger sample. Large effects may be viewed as preliminary evidence of relationships among variables; however, these findings would benefit from further investigation either via replication or qualitative methods. **RQ3.** *Propensity Score Matching/Quasi-Experimental Research.* It should be noted that while propensity score matching was used to create comparison groups that were similar to the families participating in MHC, the process cannot control all bias and should not be considered equivalent to a true experimental study. The analyses may be limited by the existence of variables that predict family participation and/or outcomes but were not available to the evaluation team. These analyses should be interpreted as only preliminary evidence of impacts (Sommers et al., 2013). *HFPI.* Overall, the HFPI appeared less sensitive to detecting changes in family outcomes over time. Currently, it is unclear if these issues are the result of the outcomes themselves (i.e., the types of outcomes the HFPI is designed to detect are particularly resistant to change) or measurement problems (i.e., issues with the scale itself or how it is administered or scored). *Fidelity.* As noted in RQ1, while fidelity has improved since the Localized MHC enhancement roll out in 2016, opportunities for sites to improve the implementation of MHC remain, especially in the areas of documentation and family review. It is important to interpret the findings within the context of Indiana's ongoing efforts to improve fidelity. Specifically, while families within the High Fidelity MHC treatment group had higher fidelity scores than families in the Low Fidelity MHC group, it is likely that many of the High Fidelity MHC families received services that did not satisfy model requirements. Moreover, because *a priori* knowledge was limited, a median split was used to

determine placement into the high and low fidelity treatment groups. Future research should identify high and low fidelity families based on the clusters of fidelity scores (those below -1, those clustering around 0, and those above 1) identified in the current study (see exploratory analyses reported in RQ3a). *Home Visitor Fidelity Measures.* Structural fidelity as measured by the home visitor survey (IN MHC Fidelity Scale) was not associated with structural fidelity as measured by the secondary activity data. Additional review of these measures suggested that home visitors may not be the best source of information for some aspects of structural fidelity. While they should have strong insights related to the model expectations in which they are directly involved (e.g., clinical consultation, reflective practice, and training), accurately recalling the extent to which all of their families were discussed in clinical consultation or how often training was offered over the course of a year might be difficult. Moreover, home visitors may be less attuned to clinical risk assignment and monthly family reviews because they are not directly involved in these processes. Finally, psychometric concerns emerged related to the IN MHC Fidelity Scale. The development of new tools to assess fidelity should be considered as part of ongoing fidelity improvement.

E. Implications of Evaluation Findings

The results of this evaluation suggest that participation in MHC may provide benefits for families, home visitors, and agencies; however, implementing the enhancement with fidelity to the model is essential for improving outcomes for stakeholders. For families in particular (RQ3), the evaluation found that increased fidelity to the treatment model was associated with improved outcomes in the areas of *emotional/verbal responsivity* ("the communicative and affective interactions between the caregiver and the child"), *organization of physical and temporal environment* ("how the child's time is organized outside the family house, [and] what the child's personal space looks like), and *parent involvement* ("how the adult interacts physically with the child") (Totsika & Sylva, 2004, p. 26). Moreover, matched-comparison analyses may suggest that when implemented with higher levels of fidelity, MHC could provide some mitigation for negative parenting outcomes experienced by families in high-risk communities. The evaluation also provided preliminary evidence to identify the model components that appear to have the strongest relationships with family outcomes when implemented with fidelity. Specifically, preliminary evidence points to possible benefits specifically associated with family reviews, reflective practice, clinical consultation, clinical risk assignment, and MHC training. Moreover, there appears to be a link between longer family participation in MHC and improved outcomes. However, it is important to note that given the small effect sizes observed for all family outcomes, the magnitude of family improvements may be very limited when fidelity is improved. Specifically, increased fidelity to the treatment accounted for approximately 1% to 2% of the variability in the outcomes. For home visitors (RQ2), the study identified how individual dimensions of MHC fidelity contribute to better outcomes. Specifically, the quality of MHC delivery was positively associated with perceived *quality of resources*. Increased home visitor responsiveness (i.e., confidence completing key aspects of MHC) was associated with greater *self-efficacy* (i.e., confidence providing support in the areas of drugs/alcohol, mental health, partner violence, behavior management, and child development) and greater *compassion satisfaction* (i.e., job satisfaction related to helping others) among the home visitors. While the effect size was small, there was some evidence to suggest that when MHC is implemented with higher levels of structural fidelity (i.e., model adherence and exposure), there was greater *burnout* among home visitors, and this finding is important to consider as adaptations to the model are considered. At the program level (RQ1), qualitative interview responses indicated that

while improvements to implementation fidelity have occurred, some sites continue to struggle with documentation and monthly reviews for all MIECHV-funded families. Moreover, at the time of the interviews, strategies to maximize fidelity were still in the early stages of implementation. Mental health consultants and HFI program managers identified a variety of factors that are currently in place that support the implementation of MHC, as well as barriers that hinder implementation with fidelity. Finally, respondents provided recommendations for improving the quality of implementation. In summary, findings examining family and home visitor outcomes (RQ3 and RQ2, respectively) provide preliminary support for continuation of the model, especially with strategies in place to improve the fidelity of implementation. Interview responses (RQ1) provide valuable insight to understanding deficiencies in current implementation, identifying strengths and gaps, and developing strategies to strengthen the model and its implementation. These data have and will continue to inform program-level decision-making in Indiana.

F. Lessons Learned

Results from the FY2018 evaluation reinforce the importance of fidelity as a necessary ingredient for improving outcomes for MHC families and home visitators. As noted in the literature, diminished fidelity of implementation can limit program outcomes, causing fruitful interventions to appear ineffective (Breitenstein et al., 2010; Fixsen et al., 2005; Mihalic, 2004). In the case of MHC, the FY2018 evaluation results have established clear links between increased fidelity and improved outcomes for both families and home visitors; however, some effects were very small. Along with promoting fidelity generally, the results also demonstrate the unique contributions of individual fidelity dimensions, which supports the development of robust strategies to improve implementation that would include a focus on delivery quality, participant responsiveness, exposure, and adherence (James Bell Associates, 2009). Moreover, resources may be maximized by prioritizing model components that are positively associated with improved stakeholder outcomes, whenever possible. Potential effects on home visitor burnout (as noted in RQ2) should be considered as strategies are developed. Mental health consultant and program manager interview responses provide a clear description of existing supports and barriers, as well as additional resources that could support improved implementation. In summary, these results provide data necessary to identify strengths and gaps in implementation, isolate critical model components to maximize resources, and develop practical strategies for improving the MHC model. Preliminary findings from the FY2018 evaluation were presented to DCS Prevention Staff during spring 2020, and these data were used to guide the development of Indiana's FY2020 MIECHV application, particularly through strategies to improve MHC and increase fidelity across sites via a variety of new supports and resources. Moving forward, the data suggest that by increasing MHC fidelity across participating sites, Indiana may experience improved outcomes for families participating in the program; however, because of the small effect sizes observed, the magnitude of family improvements may be limited as fidelity is improved.

VII. Evaluation Design

A. Entities/Organizations Responsible for Collection and Reporting Evaluation Data

Diehl Consulting Group (DCG) is an Indiana-based evaluation firm with offices in Evansville and Indianapolis. Dr. Dan Diehl and Sam Crecelius were Co-Principal Investigators for the FY2018 evaluation and were supported by Dr. Kelly Goedert, field consultant, and Amanda Vote

and Jennifer Bellville who are consultants in the group. In partnership with Indiana Department of Child Services (DCS), DCG 1) designed/identified survey instruments, 2) managed data collection, 3) cleaned and analyzed data, and 4) reported evaluation findings and recommendations.

B. Evaluation Rationale

The FY2018 evaluation was designed to 1) identify the supporting factors and barriers associated with implementing MHC with fidelity at the site level, 2) examine the effects of perceived MHC fidelity on staff perceptions of quality/relevance/usefulness of resources, self-efficacy, secondary trauma, compassion satisfaction, and burnout, and 3) explore the effects of MHC services on parenting and family functioning/support using a dose-response framework that considered increased fidelity to the treatment model as an increased dose.

RQ1. The evaluation examined the supporting factors and barriers associated with implementing MHC with fidelity. While not an evidence-based model, specific model expectations have been developed and are governed by contractual obligations agreed to by sites and staff. Concerns related to implementation fidelity emerged during the FY2016 evaluation, and it appears that some families and home visitors did not receive the level of service dictated by the MHC model. Program sites varied in the extent to which they offered the enhancement with fidelity. Based on these findings, associated evaluation recommendations, and supported by the implementation-focused evaluation approach (Stetler et al., 2006) and the Consolidated Framework for Implementation Research (CFIR) framework, Indiana utilized the FY2018 evaluation to identify 1) discrepancies that existed between the model expectations and implementation, 2) resources to support improved Mental Health Consultation implementation fidelity, and 3) actionable barriers associated with implementing the MHC model with fidelity. To support this work, the FY2018 evaluation incorporated a qualitative study designed to examine supporting factors and barriers to implementing MHC with fidelity, as reported by site-level staff.

RQ2. The study examined the extent to which home visitors' perceptions of MHC fidelity were associated with ratings of quality/relevance/usefulness of resources, self-efficacy, secondary trauma, compassion satisfaction, and burnout. The current study draws from recent studies of reflective practice by focusing more intentionally on staff outcomes that are aligned with MHC, as opposed to targeting only general home visiting outcomes. Specifically, staff outcomes selected for the FY2018 evaluation emerged from FY2016 evaluation findings (e.g., were significantly associated with MHC participation based on home visitor surveys or emerged as key themes in home visitor interview responses). Based on these findings, FY2018 staff outcomes included perceived quality/relevance/usefulness of resources, self-efficacy, secondary trauma, compassion satisfaction, and burnout.

RQ3. The study examined the relationship between participation in the MHC model and family outcomes using a dose-response framework that considered increased fidelity to the treatment model as an increased dose. Implementation fidelity was identified as a concern during the FY2016 evaluation. The FY2018 evaluation incorporated fidelity into the examination of family outcomes via exploratory analyses designed to 1) determine which fidelity measures were related to better family outcomes, 2) examine whether overall fidelity score predicted family outcomes, and 3) compare the outcome trajectory of high fidelity and low fidelity MHC participants to propensity matched non-MHC families. The selection of family outcomes was focused solely on domains that are theorized to be influenced by the MHC model. Furthermore, the FY2018

evaluation employed Mixed Linear Modeling (MLM) to increase the rigor of the study and reduce limitations found in prior studies.

C. Description of Adaptation/Enhancement

HFI MHC services are currently provided to all MIECHV-funded families receiving HFI services. This enhancement was originally conceived as a centralized provision of service through a single contractor. In January 2016, MHC transitioned to a localized provision of service with the intention of increasing availability of the consultant to direct program staff and the cultural responsiveness of the consultant to the local needs. Services are provided by licensed mental health clinicians who are located within the local agency. In each site, a locally-based, licensed clinician is hired or contracted as the mental health consultant and is available to support home visitors between two and five days per week. Core responsibilities include: (a) monitoring all MIECHV-funded family records (newly enrolled and on-going families) to assess risk and identify families with greater need; (b) one hour per month of reflective practice with home visitors serving MIECHV-funded families; (c) reviewing cases with home visitors and assisting home visitors in developing strategies to address client mental health; and (d) monitoring and identifying overall trends related to mental health concerns in program sites and conducting related trainings.

Table 2. Comparison of model components

Original Model	Revised Model (2016)
Single licensed clinician provided oversight for three Advanced Family Support Specialists (AFSS).	No "oversight" of mental health consultants as a group beyond contract requirements. Reflective supervision group available for clinical support.
Licensed clinicians did not meet directly with home visitors.	Licensed clinician(s) meets directly with home visitor.
AFSS met with home visitors.	AFSS role eliminated; services provided directly by licensed clinician(s).
Three AFSS served nine sites in seven counties.	Localized licensed clinician(s) serving each site.
AFSS may not be from or familiar with local community.	Licensed clinician(s) from local community.
AFSS reviewed family record.	Licensed clinician(s) reviews family record.
AFSS provided MHC.	Licensed clinician(s) provides MHC.
AFSS provided reflective supervision as part of the HFI model.	Licensed clinician(s) provides reflective practice in addition to HFI model supervision.
Reflective supervision counted as standard HFI model supervision.	Reflective practice and consultation beyond standard HFI model supervision.
HFI model supervisor was not present for consultation or reflective supervision.	HFI model supervisor is usually present for consultation and reflective practice ensuring continuity of guidance for home visitor.
AFSS were based at centralized locations, not at site.	Licensed clinician(s) based at local site.
AFSS provided training upon request.	Licensed clinician(s) required to provide staff training at a minimum of once every other month.

D. Use of Prior Evaluation Findings

Prior evaluations of MHC have reported mixed findings when examining the relationships between the enhancement and family and staff outcomes. Moreover, concerns related to implementation fidelity have emerged.

FY2011. The FY2011 evaluation found: 1) participation in the MHC model did not result in improved outcomes for MIECHV-funded mothers when compared to non-MIECHV-funded mothers; 2) MIECHV-funded families in sites receiving MHC had a smaller percentage of completed home visits than families in sites not receiving MHC; 3) home visitors receiving consultation reported tendencies to agree that they had increased access to professionals other than their supervisor, but they were less likely to exhibit a tendency to agree that they were

provided tools and strategies to address mental health, substance use, and domestic violence; and 4) home visitors receiving MHC were more likely to remain in their role as home visitors, even after controlling for years of experience and quarterly caseloads, than home visitors who did not receive consultation.

FY2015. The FY2015 evaluation found: 1) home visitors and other site-level staffs' perceptions of overall satisfaction with and effectiveness of the HFI MHC model suggested that MHC is supporting and building the capacity of home visitors; 2) despite reporting positive experiences with MHC, home visitors did not report levels of burnout that were statistically different from home visitors in comparison sites, and home visitors did not demonstrate high levels of job retention because of receiving consultation; and 3) participation in the MHC model did not result in improved outcomes for MIECHV-funded families.

FY2016. The FY2016 evaluation found: 1) home visitors participating in the enhancement experienced job-related outcomes similarly to non-participants, with the exception of access to resources; 2) the relationship between MHC and retention appears to be indirect, with home visitors perceiving it as a resource for stress and burnout; and 3) program sites varied in the extent to which they offered the enhancement with fidelity; however, overall, home visitor and mental health consultants provided evidence of quality when describing MHC. Moreover, there was evidence that home visitors perceive MHC as relevant and/or useful.

FY2018. The FY2018 evaluation built on prior evaluations by including new research questions that were designed to utilize previous findings and increase rigor. Concerns related to implementation fidelity emerged during prior evaluations (with specific component-level concerns identified in FY2016). Based on these findings and associated evaluation recommendations, Indiana explored site-level model implementation to inform additional implementation supports and/or adaptations to the MHC model. To explore model implementation further, the FY2018 study incorporated a qualitative study designed to examine supporting factors and barriers to implementing MHC with fidelity, as reported by site-level staff. Secondly, the FY2018 study continued prior work to explore potential benefits to home visitors who participate in the enhancement. Specifically, outcomes were selected based on promising findings from FY2016 related to access to resources, which were identified in home visitor surveys and interview responses. Moreover, the target population included only home visitors participating in the enhancement, and staff outcomes were compared by home visitors' ratings of MHC fidelity, approaches unique to the FY2018 evaluation. This approach was informed by the FY2016 fidelity study, as well as prior findings related to the impact of consulting relationships on staff outcomes (e.g., Green et al, 2004). Finally, the FY2018 evaluation included an exploratory component to examine the relationship between the enhancement and family outcomes from the HFPI and the HOME. While these outcomes were explored in prior studies, the current study differed in four distinct ways. The FY2018 evaluation explored family outcomes using a dose-response framework that considered increased fidelity to the treatment model as an increased dose. This approach incorporated fidelity into the examination of family outcomes via exploratory analyses designed to 1) determine which fidelity measures are related to better family outcomes, 2) examine whether overall fidelity score predicts family outcomes, and 3) compare the outcome trajectory of high fidelity and low fidelity MHC participants to propensity matched non-MHC families. Secondly, the FY2018 evaluation

employed Mixed Linear Modeling (MLM) to increase the rigor of the analytic approach. In prior evaluations, the potential effects of site-level and home-visitor-level differences on family outcomes were unknown, creating a significant limitation for these studies. The MLM approach helped to address limitations identified during prior studies by accounting for the hierarchical (or nested) nature of the data in the analysis. Third, whereas prior studies used statistical controls to account for implementation of a new model in January 2016, the FY2018 study focused only on the localized MHC model (described above). Finally, the study only utilized subscales from HFPI and HOME that are directly aligned with the theory of change and/or have demonstrated promising findings in prior evaluations (including non-MIECHV-funded research).

E. Theory of Change and Evaluation Framework

Theory of Change. The MHC model was designed to support home visitors, strengthen home visitors' skills, and increase home visitors' ability to help at-risk families. As recommended by Segal, Opie, and Dalziel (2012), the MHC enhancement is theory-driven and incorporates a defined theory of change. MHC is based on the theory that change occurs within the context of the relationships that consultants build with home visitors. Specifically, MHC is theorized to improve service quality by providing individualized resources and guidance, modeling reflective practice, and supporting staff whose work is emotionally challenging (Hunter, Davis, Perry, & Jones, 2016; Watson, Bailey, & Storm, 2016). The consultant establishes a climate of mutuality, reciprocity, and collaboration through which he or she moves beyond solely providing one-way instruction to promoting exploration, modeling relationships, and potentially altering the home visitor's internal experience (Johnston & Brinamen, 2012). Through parallel process, home visitors replicate these relationships in their work with families using reflective practice. Moreover, reflective practice provides support for practitioners working in emotionally-charged, stressful situations (Johnston & Brinamen, 2012; Watson, et al., 2016). Positive shifts in the family-home visitor relationship; improved self-efficacy, and capacity for reflection; and the additional resources provided by the mental health consultant are theorized to improve the quality of home visiting services (Hunter et al., 2016; Johnston & Brinamen, 2012; Watson et al., 2016).

Evaluation Framework. The FY2018 evaluation utilized a mixed-methods approach to answer the research questions, and in doing so, incorporated several evaluation approaches. Specifically, the program incorporated formative evaluation approaches (process evaluation, implementation evaluation) to address research question 1 and summative evaluation approaches (outcome-based evaluation) to examine research question 2 and research question 3 (Church & Rogers, 2006; Trochim, 2000). These approaches provide the most effective vehicles through which evaluation questions drawn from the theory of change could be answered. The formative evaluation utilized for research question 1 was supported by the implementation-focused evaluation approach (Stetler et al., 2006), which includes a focus on "resolving actionable barriers, identifying levers of change, and refining components of the implementation" (CFIR, 2019, np). Moreover, the evaluation incorporated principles of the utilization-focused evaluation approach to maximize its usefulness for key stakeholders (Patton, 2012). Through this approach, key stakeholders were fully engaged in the planning process and the evaluation activities.

F. Outcomes

Table 3. Summary of evaluation outcomes by research question.

Research Question	Outcomes	Data Collection	Analysis

RQ1. What are the supporting factors and barriers associated with implementing MHC with fidelity to the MHC model?	1) Perceptions of MHC model implementation (e.g., feasibility, supporting factors, barriers)	Program Manager Interviews: *Program Manager Semi-Structured Interview Guide* Mental Health Consultant Interviews: *Mental Health Consultant Semi-Structured Interview Guide*	Framework Analysis
RQ2. To what extent are home visitor perceptions of MHC fidelity associated with ratings of perceived quality/relevance/usefulness of resources, self-efficacy, secondary trauma, compassion satisfaction, and burnout?	Short-term Outcomes: 1) Home visitor self-efficacy, 2) quality/relevance/usefulness of resources, 3) home visitor burnout, 4) home visitor compassion satisfaction, and 5) home visitor secondary trauma	Home Visitor Surveys: (1) *Professional Quality of Life Scale*, (2) *IN MIECHV Survey for HFI Home Visitors,* (3) *IN MHC Resources Scale*	Partial-Least-Squares Path Analysis
RQ3. What is the effect of MHC on parenting and family functioning/support outcomes?	Intermediate Outcomes: 1) depression, 2) personal care, 3) mobilizing resources, 4) emotional and verbal responsivity of primary caregiver, 5) organization of physical and temporal environment, and 6) parent involvement with child	Family Outcome Data: (1) *Healthy Families Parenting Inventory* (HFPI) and (2) *Home Observation for Measurement of the Environment Inventory* (HOME)	Mixed Linear Modeling

G. Target Populations

RQ1. The target population consisted of all program managers and mental health consultants employed by HFI sites serving MIECHV-funded families (using the MIECHV MHC model). The interviews focused only on current consultants and managers. Consultants and managers providing the enhancement for fewer than three months were excluded from the sample. HFI program managers and mental health consultants were selected to participate in evaluation activities identified under RQ1 because of their firsthand experience with the implementation of MHC model components.

RQ2. Responses were drawn from the full population of home visitors who received MHC at HFI sites serving MIECHV-funded families (using the MHC model). Home visitors were selected to participate in evaluation activities identified under RQ2 because of their firsthand experience with the staff outcomes explored by this research question.

RQ3. The matched comparison groups were drawn from the full population of families participating in the HFI program from January 1, 2016 to December 30, 2019 (N = 19,302 families) for whom outcome data were available (N = 7,943 families). This consisted of 1,692 MIECHV-funded families receiving the MHC enhancement and 6,251 non-MIECHV-funded families. Families who enrolled in HFI prior to implementation of the Localized MHC model (i.e., prior to January 1, 2016) were excluded from the sample. To limit contamination, non-MIECHV-funded families (not receiving MHC) in sites serving MIECHV-funded families (receiving MHC) were excluded from the comparison groups. Families in Decatur County were excluded from the comparison group because they implement a different model of MHC through a separate funding source. Similarly, the HFI site in Marion County that does not serve MIECHV-funded families was excluded because they implemented a prior model of MHC.

H. Evaluation Questions

To address study goals, three research questions were identified.

Table 4. Summary of research questions.

Research Question 1. What are the supporting factors and barriers associated with implementing Mental Health Consultation with fidelity to the Mental Health Consultation model?

RQ1a. What discrepancies exist between the model expectations and implementation at the site level?
RQ1b. What resources would support improved Mental Health Consultation implementation?
RQ1c. What actionable barriers impede Mental Health Consultation implementation?

Research Question 2. To what extent are home visitor perceptions of Mental Health Consultation fidelity (as measured by the *IN MHC Fidelity Scale*, *Reflective Supervision Rating Scale*, and *Reflective Supervision Self-Efficacy Scale*) associated with ratings of perceived quality/relevance/usefulness of resources, self-efficacy, secondary trauma, compassion satisfaction, and burnout as measured by the *IN MIECHV Survey for HFI Home Visitors, IN MHC Resources Scale*, and the *Professional Quality of Life Scale*?

Research Question 3. What is the effect of Home Visitor Mental Health Consultation on parenting and family functioning/support outcomes as measured by the depression, personal care, and mobilizing resources subscales of the *Healthy Families Parenting Inventory* (HFPI) and the emotional and verbal responsivity of primary caregiver, organization of physical and temporal environment, and parent involvement with child subscales from the *Home Observation for Measurement of the Environment Inventory* (HOME)? *(i.e., 1) Among the families with Mental Health Consultation, do those with higher fidelity to the MHC model have better outcomes; and 2) Do families receiving MHC have better outcomes than families not receiving the enhancement?)*

RQ3a. Does fidelity to the MHC treatment model predict family outcomes within the MHC treatment group?
RQ3b. Do high and low fidelity MHC treatment groups have better family outcomes than non-MHC comparison groups?

I. Evaluation Design

Table 5. Summary of evaluation design.

Research Question	Design	Aims	Measurement
RQ1	An implementation-focused evaluation (CFIR, 2019; Stetler et al., 2006) utilizing a qualitative design and semi-structured interviews was employed.	The study explored staff perceptions of model implementation. Interview responses were drawn from the full population of 1) HFI Program Managers and 2) Mental Health Consultants.	Responses were collected through semi-structured interviews using interview guides.
RQ2	An exploratory, correlational design was employed.	The study examined the relationship between home visitors' ratings of MHC fidelity and ratings of perceived quality/relevance/usefulness of resources, self-efficacy, secondary trauma, compassion satisfaction, and burnout.	Ratings were drawn from survey responses provided by home visitors participating in the enhancement.
RQ3	An exploratory mixed methods design was employed that utilized dose-response correlational and matched-groups designs.	The FY2018 evaluation incorporated fidelity into the examination of family outcomes via exploratory analyses designed to 1) determine which fidelity measures were related to better family outcomes, 2) examine whether overall fidelity score predicts family outcomes, and 3) compare the outcome trajectory of high fidelity and low fidelity MHC participants to propensity matched non-MHC families.	**Family outcomes** were measured using the Healthy Families Parenting Inventory (HFPI) and Home Observation for Measurement of the Environment Inventory (HOME). **Fidelity measures** included the IN MHC Fidelity Scale, Reflective Supervision Rating Scale, Reflective Supervision Rating Scale – Adapted for Mental Health Consultant, Reflective Supervision Self Efficacy Scale, and Secondary Activity Reports.

J. Rationale for Design

RQ1. A qualitative design using semi-structured interviews was selected because interviews allowed the interviewer to examine specific pre-determined concepts, while providing the flexibility to explore particular themes or concepts in greater detail. The use of structured interviews utilizing an interview protocol (as opposed to an interview guide) was considered, but because of the exploratory nature of the study, final questions were open-ended and included prompts to encourage new ideas.

RQ2. The FY2016 evaluation suggested that MHC varied in terms of model adherence, exposure, quality, and participant responsiveness across sites. Moreover, prior studies have not revealed relationships between MHC participation and home visitor outcomes that are consistent with the theory of change and logic model. The literature suggests that fidelity concerns may explain why interventions that performed well in research environments fail to achieve desired results when employed in real-life conditions (Breitenstein et al., 2010; Fixsen et al., 2005; Mihalic, 2004). Based on this, a correlational design was selected to explore the association between consultation fidelity and staff outcomes. Due to the nature of the intervention, it was not feasible nor appropriate to manipulate the fidelity of MHC implementation (e.g., ethical considerations, inability to randomize subjects and locations) (Harris et al., 2006). Given the exploratory nature of the research question, the current design provided the opportunity to examine the variables of interest in their natural setting and support future research.

RQ3. The study used a multi-pronged approach to examine the relationship between participation in MHC and outcomes for families, while examining the role of implementation fidelity. Specifically, the study explored both the dose-response relationship between treatment and family outcomes (with increased fidelity of implementation defining an increased dose) and compared outcomes between treatment (high fidelity and low fidelity) and matched non-treatment groups. Along with responding to fidelity concerns identified in the FY2016 evaluation, research question 3 was designed to address evaluation technical assistance recommendations provided to Indiana during 2017. Several analytic revisions, including MLM, were made to increase study rigor. Moreover, the selection of family outcomes was guided by HomVee guidelines (Avellar and Paulsell, 2011) and focused solely on domains that were theorized to be influenced by the MHC model.

K. Evaluation Timeline

Table 6. Evaluation Timeline

Deliverable Task	2018			2019												2020									
	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct
Initial Evaluation Planning	■	■	■	■																					
Ongoing Evaluation Planning and Communication					■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■
Submit Initial Evaluation Plan				■																					
DOHVE, HRSA OPRE Review				■	■	■	■	■																	
Evaluation Plan Revisions						■																			
IRB Submission								■																	
Family Data Collection (RQ3)*	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■										
Secondary Activity Data (RQ3)*	■	■	■	■	■	■	■	■	■	■	■	■	■	■	■										
Initial Propensity Score Model											■	■													
Initial Family Data Export/Exploration (RQ3)									■	■															
Staff Survey (RQ2/RQ3)									■	■	■	■													
Mental Health Consultant Survey									■	■	■	■													
Semi-Structured Staff Interviews (RQ1)												■	■	■	■	■									
Final Propensity Score Model													■	■											
Final Family Data Export (RQ3)													■	■											
Preliminary Analyses															■	■	■								
Preliminary Report																		■	■	■					
Final Analyses																		■	■	■	■	■	■	■	■
Final Report																								■	■

*Note: Administrative data collected as part of normal service delivery and independent of the FY2016 evaluation.

L. Instruments

RQ1. Semi-Structured Interview Guide. Semi-structured interview guides (specific instruments for each population) were developed for the current project. The interview guide provided the interview questions, prompts, and guidance for the interviewer regarding the structure of the interview (Cherry, 2000; Fowler, 2004; Garvin, Cannuscio, & Branas, 2013; Lindof & Taylor, 2011). Open-ended interview questions were developed to determine the factors influencing successful implementation. The FY2016 evaluation found that HFI sites did not meet MHC model expectations for direct family services (e.g., clinical risk assignment, monthly review) and home visitor services (e.g., clinical consultation, reflective practice, training). Based on these findings, the items developed for the interview guide include a specific focus on the expectations that were not met (as reported in the FY2016 evaluation), along with a general focus on implementation fidelity. Through the CFIR-recommended implementation-focused evaluation approach, the evaluation focused on identifying discrepancies that existed between the MHC model expectations and implementation, resources to support improved Mental Health Consultation implementation, and actionable barriers.

RQ2. The Professional Quality of Life Scale. The 30-item *Professional Quality of Life Scale* (ProQOL) (Stamm, 2009) was designed to assess compassion satisfaction and compassion fatigue among professionals in the helping professions. Compassion satisfaction was measured by the *Compassion Satisfaction Subscale* and compassion fatigue was measured by the *Burnout* and *Secondary Traumatic Stress Subscales*. Items were scored on a 5-point Likert-type scale (Never to Very Often). Across multiple studies, acceptable levels of internal consistency have been observed for each of the subscales: *Compassion Satisfaction* (α = .77 to .88), *Burnout* (α = .54 to .82), and *Secondary Traumatic Stress* (α = .69 to .81) (Galiana et al., 2017; Samson et al., 2016; Stamm, 2010). In the current study, moderate levels of internal consistency were observed for *Compassion Satisfaction* (α = .91), *Burnout* (α = .77), and *Secondary Traumatic Stress* (α = .76). *Rationale:* MHC is theorized to improve service quality by providing support for practitioners working in emotionally-charged, stressful situations (Johnston & Brinamen, 2012; Watson, et al., 2016). The selection of this outcome was based on this theory and reflected in the logic model.

RQ2. IN MIECHV Survey for HFI Home Visitors. The *IN MIECHV Survey for HFI Home Visitors* utilizes the *Self-Efficacy* subscale (6 items) from the *MIHOPE Home Visitor Survey – Baseline* that was modified for the purposes of this study. The original survey was developed as part of the Mother and Infant Home Visiting Program Evaluation (MIHOPE), a large-scale evaluation of home visiting programs receiving MIECHV funding (Michalopoulos et al., 2013; 2015). Each item was scored on a 7-point Likert scale (Strongly Agree to Strongly Disagree). In the current study, acceptable levels of internal consistency were observed for the *Self-Efficacy* scale (α =.89). *Rationale:* Through MHC, mental health consultants provide individualized resources and guidance, model reflective practice, and support staff whose work is emotionally challenging. As such, staff who are provided high quality MHC are theorized to have higher levels of self-efficacy.

RQ2. IN MHC Resources Scale. The 11-item IN MHC Resources scale was developed for the current study based on findings from the FY2016 study. Each item was scored on a 5-point Likert scale (Strongly Agree to Strongly Disagree). The study examined psychometric properties using Scaling Procedures in SPSS (Green & Salkind, 2011) and found no problematic items. A principal components analysis yielded one factor that explained 85% of the variance. All items loaded successfully on the first factor, with loadings of .79 or greater. These findings were

consistent with intentions of the scale developers. In the current study, high levels of internal consistency were observed for the scale (α =.98). *Rationale:* The scale was developed to examine the quality, relevance, and usefulness of resources provided through MHC. Items were informed by home visitor and mental health consultant interview responses from the FY2016 study.

RQ2/RQ3. IN MHC Fidelity Scale. The 11-item IN MHC Fidelity Scale was designed for the current study to measure adherence and exposure to the MHC model based on Indiana's specific model expectations (as defined by MHC contractual requirements). Two subscales were developed: *Model Adherence* (4-items) and *Model Exposure* (7-items). *Model Adherence* was scored on a 4-point Likert-type scale (No Families to All Families), and *Model Exposure* was scored on a 7-point Likert-type scale (Never to More Than Once a Month). Because the IN MHC Fidelity Scale was developed for the current evaluation, psychometric properties were examined. A confirmatory principal components analysis of the four *Model Adherence* items and seven *Model Exposure* items did not confirm the structure of the scale as designed. All of the *Model Adherence* items and items 1-4, 7 from the *Model Exposure* scale loaded together onto a single component. A second component was comprised of *Model Exposure* items 5 and 6. Furthermore, while the *Model Adherence* scale, as originally structured, had adequate reliability (α = .85), the *Model Exposure* scale did not (α = .75). Because of these properties, the *Model Adherence* and *Model Exposure* items were considered a single factor for the structural equation modeling. *Rationale:* This scale was developed to measure two dimensions of fidelity not addressed adequately by existing scales: adherence and exposure (James Bell Associates, 2009).

RQ2/RQ3. Reflective Supervision Rating Scale. The 17-item *Reflective Supervision Rating Scale* (RSRS) (Gallen et al., 2016) was designed to measure reflective supervision fidelity and the quality of reflective sessions as part of the Colorado Kids Connect project. Four subscales comprise the RSRS: *Reflective Process and Skills* (6 items), *Mentoring* (6 items), *Supervision Structure* (3 items), and *Mentalization* (5 items). In the current study, high levels of internal consistency were observed for the full scale (α =.87) and the subscales: *Reflective Process and Skills* (α = .91), *Mentoring* (α = .94), *Mentalization* (α = .89) and *Supervision Structure* (α = .87). *Rationale:* The scale was designed to assess the quality of reflective supervision and was designed specifically to support fidelity measurement (Gallen et al., 2016).

RQ2/RQ3. Reflective Supervision Self-Efficacy Scale for Supervisees. The 15-item *Reflective Supervision Self-Efficacy Scale for Supervisees* (Shea, Goldberg, & Weatherston, 2012) was designed to assess supervisees' confidence in participating in components of reflective supervision (e.g., "discuss observations of infants/toddlers in a way that is attentive to health, social, emotional, and cognitive capacities," "consult with my MIECHV mental health consultant to understand my own capacities and needs"). It was originally developed for an evaluation of a reflective supervision training program implemented as part of a similar initiative in Michigan. A preliminary study suggested high levels of internal consistency with a Cronbach's α of .93. Moreover, the scale was reviewed by a panel of infant mental health experts for face and content validity (Shea, Goldberg, & Weatherston, 2016). Each item was scored on a 5-point Likert-type scale (No Confidence to Extremely High Confidence). During the current evaluation, high levels of internal consistency were observed (α = .93). *Rationale:* The scale includes items that are aligned with the style of supervision provided to home visitors through the enhancement. Specifically, the scale examined the extent to which home visitors were confident in their ability to successfully complete best practice components of reflective practice.

RQ3. Reflective Supervision Rating Scale – Adapted for Mental Health Consultant. The Gallen et al. (2016) *Reflective Supervision Rating Scale* (see description above) was adapted for

the FY2018 evaluation to be completed by mental health consultants. Given the adaptation for clinicians, the study examined psychometric properties using Scaling Procedures in SPSS (Green & Salkind, 2011) and found no problematic items. In the current study, acceptable levels of internal consistency were observed for the full scale (α =.81). Due to the small sample size (*N* = 8), appropriate diagnostics were limited. *Rationale:* The scale was designed to assess the quality of reflective supervision and was designed specifically to support fidelity measurement (Gallen et al., 2016).

RQ3. Secondary Activity Reports. Data related to program outputs were drawn from mental health consultants' Secondary Activity Reports. Secondary activity data were recorded by mental health consultants as part of their normal duties. Specifically, DCS/HFI requires that mental health consultants document all consultation with home visitors in their supervision records and provide service documentation to HFI management and DCS using Secondary Activity Reports, which included families reviewed (by family number), home visitors receiving consultation, trainings conducted, and home visits shadowed. These data were entered into the HFI data system by mental health consultants using procedures established by DCS/HFI and the database provider. De-identified, limited datasets were provided to the evaluation team by the DCS Prevention Team staff. *Rationale:* Secondary Activity Reports were utilized because they provided objective documentation of the completion of model expectations.

RQ3. Home Observation for Measurement of the Environment Inventory (HOME). The 45-item HOME (Caldwell & Bradley, 1984) was designed to provide a descriptive profile of a child's experiences in his or her home environment. Specifically, the inventory assesses the presence of emotional support and cognitive stimulation in the home through six subscales: *Emotional and Verbal Responsivity of Primary Caregiver*, *Avoidance of Restriction and Punishment*, *Organization of Physical and Temporal Environment*, *Provision of Appropriate Play Materials*, *Parent Involvement with Child*, and *Opportunities for Variety in Daily Stimulation*. Statements that make up the inventory are scored "Yes" or "No" by a trained assessor following direct observation and a semi-structured interview with the primary caregiver (Totsika & Sylva, 2004). High levels of inter-rater reliability (> .80) have been reported for the HOME; however, some issues with internal consistency have been reported, with Cronbach's α coefficients ranging from .30 to .80 (Bradley, 1993; Totsika & Sylva, 2004). Due to the manner in which data were provided, item-level survey responses were not available for psychometric testing. The FY2018 evaluation focused on three subscales: *Emotional and Verbal Responsivity of Primary Caregiver*, *Organization of Physical and Temporal Environment*, and *Parent Involvement with Child*. *Rationale:* The use of the HOME was informed by the MHC theory of change and supported by prior results.

RQ3. Healthy Families Parenting Inventory (HFPI). The 63-item HFPI (LeRoy & Milligan 2004; LeCroy, Krysik, & Milligan, 2007) was designed to assess change in eight parenting domains: *Problem Solving, Depression, Personal Care, Mobilizing Resources, Role Satisfaction, Parent/Child Interaction, Home Environment*, and *Parenting Efficacy*. These nine domains were supported by an exploratory factor analysis, and each subscale had acceptable levels of internal consistency, with Cronbach's α coefficients ranging from .76 to .92 (Krysik & LeCroy, 2012). Due to the manner in which data were provided, item-level survey responses were not available for psychometric testing. The FY2018 evaluation focused on three subscales: *Mobilizing Resources*, *Personal Care*, and *Depression*. *Rationale:* The use of the HPFI was informed by the MHC theory of change and supported by prior results.

M. Design Specific Components

Description of Qualitative Approach. The content analysis approach was employed in the current study by utilizing the framework method (Ritchie & Spencer, 1994), which can be applied to either deductive or inductive research across various epistemological, philosophical, or theoretical approaches (Gale, Heath, Cameron, Rashid & Redwood, 2013; Hsieh & Shannon, 2005, Pope, Ziebland, & Mays, 2000; Ritchie & Spencer, 1994). See *X. Analytic Methods* for a detailed discussion of the analysis.

Definition and Collection of Fidelity Components. The assessment of fidelity was designed to examine four dimensions of fidelity: Adherence, Exposure, Quality of Delivery, and Participant Responsiveness (James Bell Associates, 2009). Specific instruments/tools were informed by model expectations and relevant literature. While not an evidence-based model, specific model expectations have been developed for MHC and are governed by contractual obligations agreed to by sites and staff. These model expectations informed measures of Adherence and Exposure, which included administrative data collected by mental health consultants and a new instrument developed specifically for the evaluation. For a more robust examination of fidelity, instruments were identified in the literature that assessed Delivery Quality and Participant Responsiveness within the context of reflective supervision. Instruments/tools are listed in the following table by fidelity dimension. For detailed descriptions, see *VII. L Instruments*.

Table 7. Instruments by Fidelity Dimension

Fidelity Dimension (James Bell Associates, 2009)	Instruments
Adherence	*IN MHC Fidelity Scale, Secondary Activity Reports*
Exposure	*IN MHC Fidelity Scale*
Quality of Delivery	*Reflective Supervision Rating Scale* (Gallen, Ash, Smith, Franco, & Wilford, 2016); *Reflective Supervision Rating Scale – Adapted for Clinician* (Gallen, Ash, Smith, Franco, & Wilford, 2016)
Participant Responsiveness	*Reflective Supervision Self-Efficacy Scale* (Shea, Goldberg, & Weatherston, 2012)

VIII. Data Collection Methods and Schedule

A. Data Collection

Table 8. Data collection summary.

Data Collection Activity	Data Collection Instrument(s) Used	Respondents	Frequency of Data Collection
Mental Health Consultant Interviews	Mental Health Consultant Semi-Structured Interview Guide	7 mental health consultants	Fall 2019
Program Manager Interviews	Program Manager Semi-Structured Interview Guide	7 program mangers	Fall 2019
Family Outcomes: Emotional and verbal responsivity of primary caregiver, organization of physical and temporal environment, and parent involvement with child	*Home Observation for Measurement of the Environment Inventory (HOME)*	Low Fidelity: 779 MIECHV-funded enhancement families and 779 non-MIECHV-funded families High Fidelity: 487 MIECHV-funded enhancement families and 487 non-MIECHV-funded families	2-month, 4-month, 12-Month HOME*
Family Outcomes: Depression, personal care, and mobilizing resources	*Healthy Families Parenting Inventory (HFPI)*	Low Fidelity: 779 MIECHV-funded enhancement families and 779 non-MIECHV-funded families High Fidelity: 487 MIECHV-funded enhancement families and 487 non-MIECHV-funded families	3-month, 6-month, 12-Month HFPI*
Fidelity: Home Visitor Surveys	*IN MHC Fidelity Scale, Reflective Supervision Rating Scale, Reflective Supervision Self-Efficacy Scale for Supervisees, Professional*	74 Home Visitors	Summer 2019

Data Collection Activity	Data Collection Instrument(s) Used	Respondents	Frequency of Data Collection
	Quality of Life Scale, IN MHC Resources Scale, IN MIECHV Survey for HFI Home Visitors		
Fidelity: Mental Health Consultant Survey	*Reflective Supervision Rating Scale – Adapted for Mental Health Consultant*	8 mental health consultants	Summer 2019
Fidelity: Secondary Activity Data	*Family - Clinical Risk Assignment, Family – Monthly Review by Mental Health Consultant, Home Visitor – Monthly Clinical Consultation, Home Visitor – Monthly Reflective Practice, Home Visitor – Monthly Meetings with Mental Health Consultant, Bi-Monthly Home Visitor Training*	8 mental health consultants	Ongoing*

Note: * Family data and secondary activity data collected outside of the evaluation as part of normal home visiting activities.

B. Design Specific Components

Description of Data Collected at Multiple Time Points. The family outcomes examined in RQ3 were collected at multiple time points. Specifically, the *Healthy Families Parenting Inventory (HFPI)* was completed by the primary caregiver as early as possible (by 3 months) and at 6 months, 12 months, 24 months, and 36 months from the date of baby's birth. The *Home Observation for Measurement of the Environment Inventory (HOME)* was completed by HFI staff through observation and documentation of the Target Child at 2 months, 4 months, 12 months, and 18 months. For both instruments, the study focused on administrations occurring during the first 12 months (i.e., Administrations occurring at 18 months, 24 months, and 36 months were excluded).

Description of Consistent Data Collection and Time Points. Family data were drawn from existing sources, which were collected by HFI staff and/or their associated contracted vendors and stored using the HFI data system. All family data collection was completed by HFI staff (e.g., Family Resource Specialists, Family Support Specialists) as part of normal service delivery. Strict guidelines are in place for the collection, entry, and storage of all required family data, and extensive training is provided to all staff. Indiana conducts quarterly data reviews, and unexpected outcomes, missing data, and other questions are identified by an external evaluation team and investigated at the program and/or local level in collaboration with the Indiana MIECHV state team. Ongoing issues are addressed individually with local staff members and are brought to the attention of state MIECHV coordinators. Local Implementing Agency program managers have continuous access to live local data and can run real-time reports. Quarterly data audits are used to identify missing data. These audits occur prior to preparing data for the Indiana quarterly data review. Trends are monitored by site-level and state-level Quality Assurance staff.

IX. Sample Size and Sampling Plan

A. Sampling Plan

No sampling was employed. See *XI.B. Sample and Comparison Group*.

B. Statistical Power

All conditions under which analyses were completed were consistent with assumptions utilized in the *a priori* power analyses. **RQ2.** The population of home visitors was smaller in FY2018 compared to prior evaluations. Because power issues were suspected based on the sample of 47 complete cases and 66 imputed cases, an alternative analytic plan (included in the approved evaluation plan) was utilized, which had smaller sample size requirements: partial-least-squares

(PLS) path analysis (PA). PLS-PA required smaller samples sizes, with simulation studies suggesting that between 48 and 50 respondents were sufficient for achieving a power of 0.80 for path coefficients as small as .397 (e.g., Chin, Marcolin, & Newsted, 2003; Kock & Hadaya, 2018). **RQ3.** Across all analyses, the evaluation came close to meeting or exceeded the sample sizes proposed in *a priori power* analyses included in the evaluation plan. For the matched comparison groups, the evaluation was just under the target sample size of 500 in each high fidelity group and well over that target sample size for the low fidelity groups. In the overall sample for steps 1, 2, and 3a, the target sample size of 1,000 was exceeded, with a total of 1,692 families. Furthermore, the results of the analyses (some of which were statistically significant for very small effects (e.g., $R^2 < .01$)) suggested more than adequate power to detect even small effects at the family-level (the level-1 variable) in the MLMs. **RQ3a.** Multiple exploratory MLMs were conducted to determine the ability of each of the measures of fidelity to predict family outcomes over the three measurement time points. *A priori* power analyses revealed that at a sample size of $N = 1{,}000$, statistical power associated with each of these MLMs, assuming an alpha of 0.05**,** was 1.00. To reduce the chance of Type I errors, the Benjamini-Hochberg (1995) correction was used on the *p*-values to control the false discovery rate. Therefore, *a priori* power goals were achieved. **RQ3b.** An *a priori* power analysis was conducted using PASS 2019. The evaluation team powered for the treatment by time interaction fixed effect (with time as a continuous variable), assuming an alpha of 0.05, three repeated time-points, and a small difference in the treatment versus comparison group slopes ($d = 0.20$), in a 3-level MLM. Resultant power varied as a function of the within-family correlation among the repeated measures at level-1, with power increasing from 0.73 to 0.76, to 0.82 with increasing correlations among the level-1 repeated measures (0.45, 0.50, and 0.60, respectively).

C. Design Specific Components

Strategies to Address the Risk of Contamination. Non-MIECHV-funded families (not receiving MHC) in sites serving MIECHV-funded families (receiving MHC) were excluded from comparison groups. The HFI site in Decatur County and the HFI site in Marion County that does not serve MIECHV-funded families were excluded because they implemented different models of MHC.

Baseline Equivalence. To achieve baseline equivalence, propensity score matching was employed (Rosenbaum & Rubin, 1983). HomVEE guidance for baseline equivalence was incorporated into all planning. Specifically, equivalence was established on characteristics measured before the program group received services, and equivalence was established on the groups used in the final analysis (HomVEE, 2014, 2017). The study also relied on the Sommers et al. (2013) recommendations for controlling bias using propensity scores in observational studies, which are based on the findings of Cook, Shadish, and Wong (2008) and Glazerman, Levy, and Meyers (2003). To control bias effectively, the evaluation team sought to identify comparison groups that contained 1) prescreened individuals with motivation and incentives (or deterrents) to participate that are similar to those of the treatment group, 2) individuals from close geographical proximity to the treatment group (e.g., regional), and 3) those who have similar pretest scores on the outcome of interest compared to the treatment group. Propensity score matching yielded balanced samples based on the covariates of interest, and multivariate and univariate tests revealed no evidence of imbalance, including on covariates identified by HomVEE (2014) for a moderate study rating (the highest possible for a matched groups design):

race/ethnicity, socioeconomic status, and baseline outcomes. See *XI.B Sample and Comparison Group* for a detailed discussion of baseline equivalence.

X. Analytic Methods

A. Analytic Methods

Table 9. Summary of analytic methods.

Research Question	Analytic Methods
RQ1	**Content Analysis.** Verbatim transcripts were written from digital interview recordings and used for the analysis. The content analysis process utilized the framework method (Ritchie & Spencer, 1994), which can be applied to either deductive or inductive research across various epistemological, philosophical, or theoretical approaches (Gale, Heath, Cameron, Rashid & Redwood, 2013; Pope, Ziebland, & Mays, 2000; Ritchie & Spencer, 1994). The technique involved five steps: familiarization, identifying a thematic framework, indexing, charting, and mapping and interpretation. During the Familiarization process, the first Co-Principal Investigator identified the key ideas and recurring concepts through immersion into the text (i.e., verbatim home visitor and mental health consultant interview transcripts). Specifically, the researcher thoroughly read and re-read interview transcripts and listened to interview recordings to become familiar with the whole dataset. Next, the researcher developed a Thematic Framework by identifying all key issues, concepts, and themes in the data to create a detailed coding index. This process was informed by a variety of factors including responses and concepts present in the data, as well as by existing theory, research questions, and study objectives. The coding index was reviewed by the second Co-Principal Investigator, with input from DCG senior staff with experience in qualitative research. During the Indexing stage, the first Co-Principal Investigator applied the thematic framework (developed in the prior step) systematically to the entire dataset by annotating the transcripts with codes from the index. Next, through the Charting process, the first Co-Principal Investigator synthesized data by arranging them according to the themes to which they relate using a framework matrix. Specifically, charting allowed data to be arranged and summarized, with each column representing a theme and each row a case. Other evaluation team members were consulted throughout this process. Finally, through the Mapping and Interpretation process, the researcher explored and described the associations between themes generated by the analysis (Gale et al., 2013; Moullin, Sabater-Hernandez, & Berimoj, 2016; Pope et al., 2000; Ritchie & Spencer, 1994). At this stage, the analysis focused on "defining concepts, mapping range and nature of phenomena, creating typologies, finding associations, providing explanations, and developing strategies" (Ritchie and Spencer, 1994, p. 186). The final stage of the analysis was completed by the first Co-Principal Investigator and reviewed by the second Co-Principal Investigator. Throughout the analysis, the full transcripts were regularly consulted to confirm participants' wording and the context of their remarks. Findings were reviewed and vetted by members of the evaluation team throughout the analysis. Emerging findings were shared with the DCS state team for validation prior to study completion.
RQ2	**Partial-Least-Squares (PLS) Path Analysis (PA).** Because power issues were suspected based on the sample size, an alternative analytic approach (included in the approved evaluation plan) was utilized, which has smaller sample size requirements: partial-least-squares (PLS) path analysis (PA). A smaller sample size (with N = 47 complete cases and N = 66 imputed cases), and multivariate non-normality were observed. PLS-PA estimated the set of direct model links in two steps. In a first step, the fidelity latent construct was constructed from the set of four fidelity indicators. In a second step, the path coefficients between the fidelity construct and each of the outcome measures were estimated. PLS-PA was performed using WarpPLS. For the measurement model (creating factors from items), the evaluation team used factor-based PLS Type PTH2, which created loadings using Djikstra's (Djikstra & Henseler, 2015) consistent PLS method combined with robust path analysis, whereby latent constructs are linear sums of their composite items. The evaluation team modeled linear relations on the direct links among the latent constructs and sampled to obtain standard errors and p-values using Kock's (2018) Stable3 resampling method. Constructs were examined for statistical suppression issues using the combined criteria of a path-correlation ratio no greater than 1.3 and a p-value for the absolute difference between the path coefficient and correlation not less than 0.05. None of the directed relationships exceeded the criteria, suggesting no problems with statistical suppression. Missing data were addressed using predictive mean matching.
RQ3	**Mixed Liner Modeling.** Both research questions entailed performing mixed linear modeling (MLM) with the HOME and HFPI as outcome variables. The evaluation team assessed change in the HFPI and HOME outcomes over time during families' first year in the program using separate MLM analyses on the total scores of each of these outcome variables. For analysis of the HFPI, the team used the assessments performed at 3 months, 6 months, and 12 months. For the HOME, the evaluators used those performed at 2 months, 4 months, and 12 months. In these models, time was treated as a continuous variable. The evaluation team estimated a group-level slope representing families' change over the first year in the program. There are distinct advantages to measuring change over time rather than solely examining 12-month outcomes. Examining change over time naturally accounts for differences in baseline status (i.e., family-level intercepts), and the evaluation team could observe whether baseline status is correlated with family-level slopes. In addition, change over time provided more power and stability for the parameter estimates in the model than does using 12-month outcomes.

The random effects structure of the MLM were explored using restricted maximum likelihood (REML) estimation (Singer & Willett, 2003). Given the structure of the study design, families are nested within HFI sites, with repeated measures of the HFPI and HOME outcome variables over time nested within families. The evaluation team used a model comparison approach for determining the random effects structure that best captures the variability in the data, comparing nested versions and building to the 3-level model that matches the data structure, with time at level 1, family at level 2, and HFI site at level 3. At a minimum, the team retained family-level intercepts, but also tested for family-level variability in slopes over time, and site-level variability in slopes over time. The best-fitting random effects structure was determined using null hypothesis significant testing via likelihood ratio tests (LRT) and the Akaike Information Criterion (AIC). After having determined the best-fitting random effects structure with REML, the evaluation team used that random effects structure in modeling the fixed effects in the MLM models described below using maximum likelihood (ML) estimation (West, Welch, & Galecki, 2007). All MLMs were executed in STATA 16.0 because of its greater flexibility and facility for more complex MLM models, including maximum likelihood estimation procedures, which were better able to accommodate missing data.

Question 3a. Does fidelity to the MHC treatment model predict family outcomes within the MHC treatment group?

The analysis for research question 3a involved three steps: 1) an initial exploratory analysis to determine which objective fidelity measures were related to better outcomes on the HOME and HFPI measures, 2) creation of a single summary fidelity score for each family, and 3) analyzing whether the fidelity score predicts family outcomes as measured by the HOME and HFPI scales.

Identification of Fidelity Measures Predicting Outcomes. The evaluation team had access to a number of objective measures of fidelity (directly aligned with MHC expectations), including Family – Clinical Risk Assignment, Family – Monthly Review by Mental Health Consultant, Home Visitor – Monthly Clinical Consultation, Home Visitor – Monthly Reflective Practice, Home Visitor – Monthly Meetings with Mental Health Consultant, and Bi-Monthly Home Visitor Training drawn from mental health consultants' Secondary Activity Reports. The team also had access to subjective measures of fidelity, including home visitor surveys (*IN MHC Fidelity Scale*, *Reflective Supervision Rating Scale*, *Reflective Supervision Self Efficacy Scale*) and a mental health consultant survey (*Reflective Supervision Rating Scale – Adapted for Mental Health Consultant*). Because the relationships between fidelity measures and family outcomes were unknown, the evaluation team performed preliminary MLM analyses to identify which of the fidelity measures predict family outcomes.

Creation of Fidelity Scores. A single summary fidelity score was created from a principle components analysis (PCA) of those fidelity measures identified as predictive of family outcomes. By producing a weighted linear combination of the observed fidelity measures, the PCA produced uncorrelated and orthogonal components. Scores on the orthogonal components of the PCA were weighted and summed to create a fidelity score for each family.

Creation of High and Low Fidelity Groups. High and low fidelity treatment groups were determined by a median split on the family-level fidelity scores. The median split was chosen for defining these groups because of the exploratory nature of this analysis, the lack of clear a priori threshold for determining high versus low fidelity scores, and the desire to preserve as much of the sample as possible, thus protecting statistical power to the extent possible.

Inferential Statistics. An MLM examined whether overall fidelity score predicted family outcomes as measured by the HOME and HFPI scales. To the extent that greater fidelity was associated with higher slopes over time on these outcomes, it suggested a causal relationship between receiving the intervention and having better outcomes.

Question 3b. Do high and low fidelity MHC treatment groups have better family outcomes than non-MHC treatment groups?

The analysis for research question 3b involved three steps: 1) the creation of propensity scores based on participant characteristics, 2) matching high and low fidelity treatment groups each with their own comparison group based on the propensity score, and 3) conducting analyses to determine differences between HFPI and HOME subscale scores between the matched groups.

Propensity score development. Propensity scores (i.e., the conditional probability of treatment assignment) were created using a logistic regression model that incorporated observable covariates or appropriate proxies (Austin, 2011; Caliendo & Kopeinig, 2008; D'Agostino, 1998; Rosenbaum & Rubin, 1983). The selection of covariates was informed by the relevant literature and theory (Austin, 2011; Caliendo & Kopeinig, 2008), institutional selection processes (i.e., the role of initial family risk status/scores in assignment/services) (Blundell, Deardeb, & Sianesi, 2005; Sianesi, 2004), and empirical methods (Caliendo & Kopeinig, 2008). These include demographics associated with MIECHV participation/outcomes (e.g., adult age, child age, education, employment status, income) (MIECHV, 2016), baseline screening instruments (e.g., HFI Eight Item Screen, Parent-Survey/Family Stress Checklist, IPV screens, EPDS), and other variables influencing participation and outcomes. To account for missing data, the missing indicator method was used to model the relationship between the pattern of missing data and propensity to participate in MHC (Rosenbaum & Rubin, 1984).

Covariates included the following: Baseline Edinburgh Postnatal Depression Scale, Baseline HFPI (Total Score), Baseline HOME (Total Score), Intimate Partner Violence Status, Parent Survey/Family Stress Checklist, Urban Influence Codes (USDA, 2013), Child Birth Status, Primary Caregiver's Age, Primary Caregiver's Language, Primary Caregiver's Racial Identification, Primary Caregiver's History of Substance Abuse, Primary Caregiver's History of Mental Illness, Primary Caregiver's History of Criminality, Primary Caregiver's Education Level, Primary Caregiver's Enrollment Status, and Income.

Matching. To balance the treatment and comparison groups, the research team utilized nearest neighbor matching (with caliper) using the R-Essentials SPSS extension (D'Agostino, 1998; Ho, Imai, King, & Stuart, 2007). This process involved matching treatment individuals with the comparison individual with the most similar propensity scores (D'Agostino, 1998; Stuart, 2010). The use of the caliper reduced the number of poor matches utilized in the dataset (Stuart, 2010). A caliper width of 0.2 of the standard deviation of the propensity score was used (Austin, 2011; Rosenbaum & Rubin, 1985; Cochran & Rubin). Unmatched cases were excluded from the analysis.

Inferential Statistics. The evaluation team performed two separate mixed linear models (MLM): one examining the difference between the high fidelity treatment group and its propensity score matched comparison group, and a second examining the difference between the low fidelity treatment group and its propensity matched comparison group. Because the primary research question was the effectiveness of the treatment (i.e., assignment to MHC vs. non-MHC), for both MLMs the team modeled the treatment by time interaction, and component main effects, as fixed effects (i.e., group-level effects) in all the models under consideration. The team was then able to ask whether this change differs for the treatment (MHC) and comparison groups. Secondary, exploratory MLM analyses examined whether change over time within the MHC treatment group varies as a function of the random effect of MHC consultant.

The MLM analytic method and random effects structure employed for both Research Questions 3a and 3b allowed for the examination of site-level effects in both analyses. As part of the random-effect structure, measurement timepoints were nested within family, which were nested within sites. This structure allowed the researchers to 1) assess whether there is significant variability among the sites and 2) to examine the site-level best linear unbiased predictors (BLUPs), which are site-level regression slopes for the outcome measures over the three measurement timepoints. Thus, the evaluation was able to examine the magnitude of the regression coefficient for each individual site.

B. Evaluation Cost

Total evaluation costs were $100,000 for a two-year period. Costs were based on a standard group rate of $100 per hour for consulting staff assigned to the project. The group rate includes employee salaries (commensurate with education and evaluation and analytic experience), employer expenses (fringe benefits, FICA (Social Security and Medicaid), IN-SUTA), and indirect costs. Group rates align with current market rates for proposed services. DCG committed two senior consultants, one statistical field consultant, and two consultants to support the project (500 hours per contract year).

C. Design Specific Components

Qualitative Data Analysis Techniques. The content analysis process utilized the framework method (Ritchie & Spencer, 1994) (see above) (Gale, Heath, Cameron, Rashid & Redwood, 2013; Pope, Ziebland, & Mays, 2000; Ritchie & Spencer, 1994). Consistent with this method, the identification of themes was informed by responses and concepts present in the data, as well as by existing theory, research questions, and study objectives. In the final step, the analysis focused on "defining concepts, mapping range and nature of phenomena, creating typologies, finding associations, providing explanations, and developing strategies" (Ritchie and Spencer, 1994, p. 186) to identify associations between themes generated by the analysis.

Qualitative Data Processing. Interviews were recorded (with participants' permission) and transcribed verbatim for analysis. Data processing was informed by the framework method, which involved five steps: familiarization, identifying a thematic framework, indexing, charting, and mapping and interpretation. Each is described in detail above.

XI. Evaluation Results

A. Results

Research Question 1

RQ1a. What discrepancies exist between the model expectations and implementation at the site level?

As part of the semi-structured interviews, program managers and mental health consultants were asked to indicate if specific MHC model expectations were met at their sites. Overall, they reported that they believed most requirements were met consistently; however, responses suggested that most sites experienced challenges meeting documentation expectations, and some sites struggled to complete monthly reviews for all MIECHV-funded families.

Table 10. Program manager and mental health consultant perceptions of model adherence

Model Expectations	Met	Not Met
All New MIECHV-Funded Families Reviewed	93% (n=13)	7% (n=1)
Family Priority Assigned for All MIECHV-Funded Families	93% (n=13)	7% (n=1)
Monthly Review for All MIECHV-Funded Families	79% (n=11)	21% (n=3)
Monthly Clinical Consultation for All Home Visitors	93% (n=13)	7% (n=1)
Monthly Reflective Practice for all Home Visitors	93% (n=13)	7% (n=1)
Bi-Monthly Training Provided for Home Visitors	93% (n=13)	7% (n=1)
Documentation Completed	43% (n=6)	57% (n=8)

Documentation. The majority of participants (57%, *n* = 8) reported that expectations for documentation have not been met consistently at their site. Mental health consultants and program mangers noted that the spring 2019 database transition, unclear MHC model expectations, and a lack of experience documenting their responsibilities and/or using the database contributed to difficulties completing documentation in adherence with expectations. Consultants were especially concerned that documentation issues reflected poorly on their job performance because model expectations that they believed that they had met were not accurately captured in the database.

"I know things were lost and overlooked in the [spring 2019] transfer [to a new HFI database]. I could see that there could be a stretch there where we did not meet our documentation requirements. For one thing, the [mental health consultant] didn't have access for three weeks. We couldn't get in the system at all. We didn't have the right clearance to see what we needed to see. There are still some areas I can't see… So, when [the mental health consultant] puts in a secondary activity, I can't see that [the mental health consultant] did that." – Program Manger

"I have learned more about things when it comes to documentation and things that I was supposed to be doing that maybe wasn't being reflected in documentation." – Mental Health Consultant

"I am still working towards it because I can't get through a month without errors." – Mental Health Consultant

"A bunch of us [mental health consultants] discussed this and realized that we were doing the work, [but] we didn't know it was required to document certain things. It took several months to get a response back from the people at DCS in regards to what exactly the expectation was. It was a little bit frustrating, to say the least. I'm going to be honest with you on that. Because all the work was being done, [but] there was no consistent communication until recently on how and what we should be documenting it." – Mental Health Consultant

Monthly Review of MIECHV-Funded Families. While the majority (79%) of mental health consultants and program managers reported that all MIECHV-funded families were being reviewed monthly at the time of the interview (August-September 2019), participants noted that in many sites, only high priority families had been reviewed and documented on a monthly basis prior to a May 2019 MHC training that clarified this model expectation. Specifically, 86% (6/7) of mental health consultants noted prior confusion about this expectation and stated that they had incorrectly completed and/or documented family reviews before May 2019. After receiving clarification, some sites (particularly sites with larger caseloads) continued to struggle to review

all families each month. As a result, some staff feared that other aspects of the model (e.g., training) may suffer due to the amount of time dedicated to reviewing families. Additionally, some staff were uncertain about what a review should entail and sought additional clarity. Finally, given that all families (regardless of priority) were reviewed monthly, the expectation that mental health consultants prioritize families for MHC was questioned.

"The (monthly family review) is my most unclear area of the expectation overall. I know that I'm supposed to do these reviews, but I think, for me, I'd feel more comfortable if I had an overall standard [that defines what should be reviewed]." – Mental Health Consultant

"Recently, in May [2019], we attended a meeting at a training regarding documentation for our mental health consultants. Through that meeting, it came out that the expectation was that all [MIECHV] families are discussed every month with the home visitor and the [mental health consultant], and that is not happening. That was new to everybody! I don't know if that continues to be the expectation, but we would not be able to meet that." – Program Manager

"We just started reviewing the entire caseloads last month… That is a concern moving forward. Something is going to have to give when you [increase the number of families reviewed by 233%]." – Mental Health Consultant*

"I've been doing the monthly review of every family since I started, but I wasn't made aware that it was something we were supposed to be writing a note in... the data system until a few months ago. That would have been the only challenge, because the work I had been doing had not been captured." – Mental Health Consultant

"It is a lot; I carry over [number redacted] MIECHV families at any given time. That is a lot of families for our [mental health consultant] to review each and every month. We are working with them to streamline, help them understand how you can review a family in a way that isn't too prohibited timewise." – Program Manager*

Note: * Caseload numbers redacted to protect participant confidentiality.

RQ1b. What resources would support improved Mental Health Consultation implementation?

Interviews were completed with seven mental health consultants and seven program managers to explore the currently available supports and additional resources needed to implement the MHC model expectations with fidelity. Using the framework method (Richie & Spencer, 1994), key themes were identified (e.g., mentioned by 40% or more of respondents) and are discussed in the following sections. Percentages are provided to present the prevalence of each theme in the responses.

Current Supports. Analysis of interview responses revealed a number of current supports, and the most common themes involved MHC alignment with existing work (93%), agency/management support for MHC (86%), collaboration among mental health consultants (57%), mental health consultants' expertise/experience (50%), and the MHC training offered in May 2019 (43%).

Alignment with Existing Work. Nearly all participants (93%) noted the value of MHC and its natural alignment with the existing work at their site (Mental Health Consultants: 100%; Program Mangers: 86%). Sites recognize the role that mental health plays in family success, as well as gaps in mental health expertise among home visitors and supervisors. Respondents noted that MHC was a necessary enhancement to support mental health needs for families, and noted that sites leverage MHC to bolster their capacity to address mental health needs and to improve family outcomes. Moreover, there is generally shared understanding among staff about the role of reflective supervision and how home visitors can use the consultant to improve their work with families, especially as it relates to mental health. From a fidelity perspective, this alignment promoted buy-in among leaders and staff. In particular, the MHC's relevance for home visitors

encouraged them to participle in the model through clinical consultation, reflective practice, and training.

"Reflective supervision is the core component of the Healthy Families model. It is something that all of our supervisors are training in doing, so this piece was not new. But I think that it is really refreshing to provide that reflection outside of the supervisory relationship, so they have someone that they can process these things with who is not also their boss." – Program Manager

"Although home visitors aren't trained or have the level of education to provide mental health support to their families, that is what comes up in almost every single visit. Whether it is old trauma, current trauma stress. Although these home visitors aren't really trained to deal with that, I think providing them that support so they can go back out to the families…know how to guide them, what to say, how to support them." – Mental Health Consultant

"I think at the current stage, it fits in amazingly. Since this project started, Healthy Families, in general has become much more aware of how mental health affects the families in the homes that they serve." – Mental Health Consultant

"It just improves overall service to families, to be able to receive that assistance in coming up with interventions." – Program Manager

Agency/Management Support. The majority of respondents (86%) described their agency and its management as a key support for implementing MHC with fidelity (Mental Health Consultants: 100%; Program Mangers: 71%). Specifically, respondents reported that management integrated MHC into all agency activities, communicated the importance of MHC to staff, ensured that time and resources were appropriately allocated for MHC, and involved mental health consultants in meetings and other collaborative work. Management buy-in reinforced model expectations to home visitors, which encouraged them to participate in consultation activities, reflective practice, and training.

"The agency is supportive of me doing those trainings, so there is some time allowed… kind of built in because it does take a lot of work to put those together." – Mental Health Consultant

"I will say my infrastructure here at our site has been very supportive of this role from the program manager, the supervisors. I think they make it very clear to staff that my role is vital to this program, and I think [management] have been very supportive on this journey with me." – Mental Health Consultant

"Everybody at the agency works pretty well together, and it is the expectation that [MHC] happens… My position is integrated into the overall process, and I think it is very supportive from the director down that [participating in MHC] is what we do, this is part of our process. That way people don't see it as another random meeting that they have to go to." – Mental Health Consultant

"A lot of times I will be in a discussion with a supervisor about a family situation, and the guidance is let's loop the [mental health consultant] in on this and get their feedback. Throughout the normal course of services as we identify issues and concerns that we think might be best supported with a clinical perspective, we are able to loop them in." – Program Manger

"I'm super supported by upper management and home visitor supervisors. They voice the benefits and the fact that MHC is trying to help home visitors in the most effective way, in their jobs, and in their own mental health. Being supported by upper management is the most helpful." – Mental Health Consultant

Collaboration Among Mental Health Consultants. Over half of all participants (57%) noted that collaboration among mental health consultants supported improved implementation fidelity. In particular, the majority (71%) of mental health consultants identified collaboration with their peers as a key support, while slightly less than half (43%) of program managers identified this support. Specifically, both one-on-one communication and the monthly reflective supervision to which all mental health consultants had access were identified as beneficial by participants. Respondents noted that they communicate informally with their peers one-on-one or in small groups on an as-needed basis through emails and telephone calls. Additionally, monthly reflective supervision was made available to all mental health consultants by DCS. Through collaboration, the consultants have shared forms for documentation, resources, strategies for

supporting home visitors, general information about MHC, and best practices for meeting model expectations.

"Those two [peer] consultants were critical for me… They have allowed me to contact them pretty freely… That was important for me, to have someone else in the same role as me, being able to share with me, not only how they did their job but what they remembered in the beginning." – Mental Health Consultant

"One of the things I appreciate the most is that we have a phone supervision that is offered twice a month with other [mental health consultants] in the state. That is the biggest supporting factor for me because I can talk to other [mental health consultants] in the state, asking how they get through certain challenges." – Mental Health Consultant

"We worked with [another site's] [mental health consultant], so I think we use similar forms." – Mental Health Consultant

Mental Health Consultants' Expertise/Experience. Half (50%) of respondents described mental health consultants' experience and expertise as key supports for implementing MHC with fidelity. This was most frequently identified as a program support by program managers (57%); however, 43% of consultants also identified experience and expertise as a support for implementation fidelity. In many cases, respondents noted that mental health consultants' backgrounds were highly aligned with the duties and tasks necessary to implement MHC model expectations with fidelity. Specifically, consultants relied on their education and clinical experience to review families, provide consultation and reflective supervision, and design and deliver training efficiently and with high quality.

"Well, the level of clinical skills that we were fortunate to have by hiring [mental health consultant] has made it easier [to implement MHC]." – Program Manger

"She definitely, with her years of being a therapist, works with staff to get to the root of how they're feeling, why they're feeling, how that could affect the services we provide, and how they could be supported through extra supervision time." – Program Manger

"[Mental health consultant] has been involved in reflective practice her entire, well, her career as a licensed clinical social worker. You can tell from her documentation that this is very natural for her to provide reflective practice." - Program Manager

"I was a LCSW for 26 years before coming into this role… I was a home-based therapist, so I had years of that experience with families who were referred from DCS or from probation. That helped me move into this role pretty seamlessly because I had seen the program from the outside, and I had worked with local mental health center and the local DCS office. I was very connected to those other resources." – Mental Health Consultant

"[Mental health consultants] use their own professional expertise to identify which families need that high-risk designation." – Program Manger

May 2019 MHC Training. The training provided to mental health consultants at the Institute for Strengthening Families in May 2019 was identified by 43% of participants as a key support for improving the fidelity of implementation (Mental Health Consultants: 43%; Program Mangers: 43%). Through this training, DCS Prevention staff reviewed the model expectations and associated documentation, answered questions, and provided resources for reference. Additionally, this training provided an opportunity for consultants to meet in-person and for them to interact with the DCS staff who oversee MHC. All mental health consultants and program managers were invited to participate.

"We had a meeting at our last Institute training in May with the managers and the MIECHV site and [mental health consultants]. I think that was really helpful for the [mental health consultants] and myself, to make sure that the expectations were clear: first, what needed to be done and secondly, how we were documenting that." – Program Manager

"Back at the beginning of the new system, there was some question about how things should get entered, but they got that taken care of at the May Institute. I think it was a little frustrating at first because it sounded like each site was doing things a little bit different, but now everyone is on the same page." – Program Manager

"They gave a very specific, extensive PowerPoint at the Institute that was very helpful. Any time I can't remember, I can pull up that PowerPoint." – Mental Health Consultant

Additional Resources/Support Needed. Along with existing resources, analysis of interview responses revealed three additional supports that respondents believed would improve implementation fidelity if provided to mental health consultants: increased opportunities for collaboration (50%), additional training that focuses specifically on MHC responsibilities (50%), and more detailed documentation guidelines and instructions (43%).

Increased Collaboration Among Mental Health Consultants. Half (50%) of participants identified increased opportunities for formal collaboration (e.g., facilitated by DCS) among mental health consultants as a strategy to improve fidelity, and this strategy was especially popular with the consultants (86%). Specifically, respondents were interested in more frequent in-person meetings and continuation of reflective supervision calls. Additionally, there was some interest in exploring scheduling options for that would increase participation in the existing reflective supervision provided to the consultants.

"That would be fabulous to interact with the other [mental health consultants] across the state. Not everybody gets on the call either, which would be nice if we all had that opportunity. Originally, they talked about bringing us together more often and that hasn't happened either. It would be helpful even if it was quarterly." – Mental Health Consultant

"Continue monthly reflective supervision that [the mental health consultant] receives, on a conference call with other [mental health consultants], and I know she enjoys that. And being able to get her peers' support on the participants in our program." – Program Manager

Dedicated Mental Health Consultant Training. Half (50%) of participants recommended dedicated training for mental health consultants that focused specifically on their role in MHC. This topic was particularly salient for mental health consultants (Mental Health Consultants: 71%; Program Mangers: 29%). Specific recommendations included reflective supervision training, funding for continuing education, and yearly sessions specifically for mental health consultants at the Institute for Strengthening Families.

"I think continued training on what reflective consultation is and how to do it well would be helpful. I'm fortunate that our full-time [mental health consultant] is amazing. She has gone out and done a lot of work on her own. We've taken advantage of some training opportunities." – Program Manger

"There is a real lack of understanding in the field of what infant mental health is and reflective consultation is because it is very specific to this early childhood/zero to three population. I think when we hire, we hire general LCSWs who do not have a background in this." – Program Manger

"I would be open and love more opportunities for that more advanced [reflective supervision] training beyond what a Healthy Families supervisor would get, really get into the deep, nuts and bolts of it. The trainings that are offered through Healthy Families are geared more towards the supervisor role, which is different than what we are doing. I think a training that is more tailored towards our role and what we are doing that dives in deep would be helpful." – Mental Health Consultant

"There needs to be some more continuous education and training for [mental health consultants] that are separate from what trainings home visitors get… I do believe providing mental health consultation, that it is always important to grow as a [mental health consultant] or as a mental health provider. So, one thing I found surprising in taking this job is that there are not monies or grants focused on making sure that the [mental health consultants] continue with their education." – Mental Health Consultant

Documentation Guidelines. Nearly half (43%) of participants noted a need for additional MHC documentation guidelines for mental health consultants (Mental Health Consultants: 43%; Program Mangers: 43%). Specific recommendations included standardized forms, additional clarification of and rationale for model expectations, documentation reviews/audits, and detailed roles and responsibilities.

"I made up my own form for doing case reviews and for doing clinical consultations and reflective practice. Sometimes I hope that I'm doing this well with having my own notes and paperwork and that they are helpful. A standard way to document would be helpful." – Mental Health Consultant

"Helping to clarify why we are designating people high-risk. What is the reason for the designation if all families get the same level of oversights? Maybe some guidance or clarity about what… record review means. Are there some key things that [mental health consultants] might want to look at?" – Program Manger

RQ1c. What actionable barriers impede Mental Health Consultation implementation? Interviews were completed with seven mental health consultants and seven program managers to explore the barriers that impede Mental Health Consultation implementation. Using the framework method (Richie & Spencer, 1994), key themes were identified (e.g., mentioned by 40% or more of respondents). Key themes are discussed in the following sections. Percentages are provided to present the prevalence of each theme in the responses.

Analysis of interview responses revealed a number of barriers, and the most common themes involved data system migration (86%), model expectation clarity (79%), caseloads (57%), mental health consultants' time capacity (57%), scheduling challenges (43%), and home visitors' buy-in (43%).

Data System Migration. Nearly all respondents identified the data system migration in early 2019 as a barrier that limited implementation fidelity (Mental Health Consultants: 86%; Program Mangers: 86%). Specific barriers associated with the migration included lack of access to the new data system (e.g., unavailable login credentials, user privileges not created, inability to edit family records), difficulty learning to use the new system and accessing technical support, limited features to support MHC (e.g., no reports that identify new families, that allow program managers to review mental health consultants' work, or that allow mental health consultants to track their own work or flag records/assessments that have already be reviewed), and poor functionality (e.g., multiple screens/menus to navigate, increased likelihood for missing data). Note: As described in *VII. K Timeline,* interviews occurred approximately six months after the initial database rollout, which may have influenced the frequency with which related barriers were described by participants.

"There was a period of time where [the mental health consultant] couldn't go in and see the families or evaluate for high-risk status." – Program Manager

"Our new database, Enlite, does not have a report that we can generate which are new families coming in. Sometimes I feel that that is an issue. When a high-risk family has been assigned outside of case assignment that I may not have been privy to, there is no monthly report I can run to see if there are any families I am missing." – Mental Health Consultant

"I can't see [the mental health consultant's] non-client secondary report. I went to the help desk, and they tried to convince me I can see it, and I cannot. For billing and to check her work and things like that, she has to send me the report." – Program Manager

"There is no way for a [mental health consultant] to check off: yes, I've read this intake. We can mark [risk status] as a yes or a no for a clinical high risk, but in order to see that, we have to click through every single assessment." – Mental Health Consultant

"The website we are using is not user friendly. Too much clicking and drop-downs, not easy to navigate" – Mental Health Consultant

"Random people are finding out ways to make the job easier, but there hasn't been any directional book or user guide… We can run reports, but they aren't easy to run. They are time consuming. We haven't figured out how to run reports better. When it is just one or two people [in the mental health consultant role] in a facility of [number redacted], nobody knows what we do, there is no one to turn to, we have to figure it out." – Mental Health Consultant

"I would assume that the risk of missing a family and not reviewing a family is much greater now than it was with the old system." – Mental Health Consultant

"We can put in a help ticket, but the last I heard is that they had hundreds and hundreds of help tickets and were working their way through. I think communication was bad at the beginning, and some of the clinicians didn't even have access, so we were having our program manager write help tickets for us to get access and there weren't any answers back." – Mental Health Consultant

Model Expectation Clarity. The majority (79%) of participants described a lack of clarity related to model expectations as negatively affecting MHC fidelity, and this barrier was particularly prevalent for mental health consultants (Mental Health Consultants: 100%; Program Mangers: 57%). Specifically, participants sought additional clarity related to family reviews and documentation expectations, which was consistent with the discrepancies identified by mental health consultants and program managers when comparing implementation at their site with model expectations.

"With the high-risk families, I don't know how clear it was made to the [mental health consultants]. They go into the database and mark if the family is high-risk or not. There are not specific criteria given to them of what makes a family high-risk. They just say it is based on our [mental health consultant's] judgment. Maybe that could be helpful if they had specific criteria for making the family high-risk." – Program Manager

"It was just in May that they told us [family review] was a monthly requirement, and you should be doing it with each worker, each month. We already kind of were, but I was not documenting it every month in the Enlite system." – Mental Health Consultant

"The question was brought up at the May meeting if we were supposed to be reviewing every [MIECHV] family or just the high-risk families. The [mental health consultants] had all been doing just the high-risk, but we were made aware in May that we needed to do every [MIECHV] family." – Mental Health Consultant

"The biggest thing that came away after the Institute this past spring was that we needed to show a consultation on every MIECHV family. That was not our understanding going into this, for several years. Our understanding was that the [mental health consultant] needed to follow the high-risk families every single month and meet with the staff once a month, but we were not under the impression that every single MIECHV family needed a consultation, meaning a record review. It was our understanding it was the families that were designated as high-risk need that monthly review. The other families, many would be discussed and reviewed through the course of reflective consultation but that it did not require that record review, every single month, for every single family." – Program Manger

Caseloads. Just over half (57%) of participants identified caseloads as barrier to implementation fidelity (Mental Health Consultants: 57%; Program Mangers: 57%). Specifically, large caseloads made it difficult for sites to meet model expectations, particularly the monthly family review and clinical consultation. In addition, larger caseloads were perceived to negatively affect the quality of MHC delivery because less time could be devoted to each family.

"We have [number redacted] families in MIECHV. There is absolutely no way [the mental health consultant] can talk about [number redacted*] families in a month when [that person] meets with a home visitor for one hour a month and especially be providing quality input." – Program Manger*

"If I have a home visitor with 10 families, then they have to talk about 10 families… What good is it doing to talk about the family for two minutes?" – Program Manger

"Having to put in [documentation] for every single person you're reading… can be time consuming. I'm already taking my own personal notes so I don't forget in my meeting with the home visitor and then having to go in to the computer and writing a separate note, it is time consuming and a I think it is a barrier." – Mental Health Consultant

"We are talking a lot more families! More than double. The families that are not high-risk, I will typically read the last home visit and any incident report. They don't get as thorough of a look over as the high-risk families. There isn't time. I'm more dependent on my supervisors to tell me if there something is up with those families." – Mental Health Consultant

Note: * Caseload numbers redacted to protect participant confidentiality.

Mental Health Consultant Capacity – Time. Approximately half (57%) of participants indicated that mental health consultants' capacity to meet expectations was limited by the amount of time that they were available to support MHC (Mental Health Consultants: 71%; Program Mangers: 43%). In some sites, staff did not believe that the amount of time that mental health consultants were on site was sufficient to address model expectations. Moreover, mental

health consultants reported that an inadequate amount of time was dedicated to clinical consultation and reflective practice, which they believed had negative effects on the quality of MHC delivery.

"[There are] still a lot of other responsibilities. [Time spent reviewing families] takes away from that one-on-one or that open availability for staff to come in to process and chat." – Program Manager

"Sometimes not having enough time during clinical supervision [is a barrier]. It's an hour, but some workers see 20-26 families. There's just not enough time to get through everybody. They already have supervision, but the mental health component of being able to help some of their families, sometimes supervision isn't long enough." – Mental Health Consultant

"I'm not convinced that one hour of reflective practice is always enough, so in that hour of reflective practice, there is not time to do that, in my opinion so we do that outside of reflective practice time. I don't know if that is what other people do. We've tried to do it within that hour but there isn't enough time. I'm not sure how it is supposed to be done, but that would be the only barrier. Making sure that we meet outside of that hour to do some of that work, which we do." – Mental Health Consultant

"I keep pretty busy since I'm only 20 hours a week." – Mental Health Consultant

Scheduling Challenges. Participants reported that coordinating schedules among mental health consultants, home visitors, and supervisors created implementation barriers. This barrier was more prevalent for program managers than mental health consultants (Mental Health Consultants: 29%; Program Mangers: 57%). Due to the nature of their work, home visitors' and supervisors' schedules are often unpredictable, and while participants noted that mental health consultants attempt to adapt to scheduling needs, canceled clinical consultation and reflective practice sessions are often difficult to reschedule.

"Every now and then, because we have to have the supervisors in there [for consultation], it can present some scheduling challenges. Because supervisors have their own supervisions they have to do every week. Because I'm trying to schedule with two people, the HV's are very busy and so are the supervisors and sometimes that can be a barrier." – Mental Health Consultant

"The problems we had were the [mental health consultant] cancelling appointments for whatever reason or the HV canceling, and it not getting rescheduled within that month. Especially when they are close to the end of the month, and there is not choice, that is one issue." – Program Manger

"Scheduling [is a barrier]. They've been doing a really good job, but you are working with HV's who are busy, in and out." – Program Manger

"Typically, when we do reflective practice with supervisors and with the home visitor, but there are times when schedules don't allow for that, so we have to be flexible and know that sometimes the home visitors see the [mental health consultant] by themselves." – Program Manger

Home Visitor Buy-In. Most program mangers identified issues with home visitor buy-in as a barrier for MHC implementation; however, this was not noted as a concern by mental health consultants (Mental Health Consultants: 14%; Program Mangers: 71%). Specifically, program managers noted that some home visitors see MHC as an extra requirement and fail to see the value in participating in the extra support.

"We have some home visitors that don't see the value in the [mental health consultant]; they feel they are already processing the family with their supervisor…and we've been working on that. They don't quite understand that the [mental health consultant] brings out another layer of support." – Program Manger

"I think, because we are a larger site, sometimes the home visitors feel they are too busy, so just reminding them that it is important to fit it in their schedule and they really aren't too busy." – Program Manger

"[Home visitors'] lives are so focused on meeting the needs of the family. That self-care piece and finding time for two supervisions [in] one week is a lot for them sometimes." – Program Manger

Research Question 2

RQ2. To what extent are home visitor perceptions of Mental Health Consultation fidelity associated with ratings of perceived quality/relevance/usefulness of resources, self-efficacy, secondary trauma, compassion satisfaction, and burnout?

To examine the relationships between home visitors' ratings of MHC fidelity and ratings of perceived quality/relevance/usefulness of resources, self-efficacy, secondary trauma, compassion satisfaction, and burnout, a partial least squares-path analysis (PLS-PA) was utilized. A fully linked model was created that explored the unique relationships between each fidelity measure and each staff outcome.

Missing Data and Imputation. Of the 77 participants starting the survey, 3 participants did not complete any of the survey questions, 47 completed all the survey questions, and 11 participants had more than 5% missing data. For the purpose of analysis, participants with more than 5% missing data were excluded. Missing data were imputed for the cases with less than 5% missing ($N = 66$) using predictive mean matching. After assuring an adequate imputation by comparing the descriptive statistics for the imputed ($N = 66$) and complete cases ($N = 47$), all subsequent analyses were performed on the imputed data. Note, none of the single survey items from the original raw dataset had greater than 5% missing cases, which suggested that no item was systematically omitted by respondents and that the assumption of missing at random appears to hold.

Descriptive Statistics. Descriptive statistics for each scale are provided below for the imputed sample.

Table 11. Descriptive statistics for fidelity and staff outcome scales.

Measures	*M*	*SD*	*N*
Staff Outcome			
Resource Quality	4.17	0.89	66
Self-Efficacy	6.03	0.86	66
Burnout	2.03	0.59	66
Compassion Satisfaction	4.20	0.58	66
Secondary Trauma	1.91	0.61	66
Fidelity Measure			
Delivery Quality	4.30	0.52	66
Participant Responsiveness	3.65	0.68	66
Structural Fidelity	4.85	0.99	66

Model Fit. Indices of model fit suggested that the fully linked model was a good fit to the data.

Table 12. Model indices of fit.

	Statistic	Interpretation
Tenenhaus Goodness of Fit (GOF)	0.304	small >= 0.1, medium >= 0.25, large >= 0.36
Average R-squared	0.187, $p = 0.028$	
Average Adjusted R-squared	0.148, $p = 0.053$	
Average Path Coefficient	0.200, p = 0.022	
Average Block Variance Inflation Factor	1.438	acceptable if <= 5, ideally <= 3.3
Average Full Collinearity Variance Inflation Factor	1.833	acceptable if <= 5, ideally <= 3.3

Path Coefficients and Characteristics of Latent Variable Composites. Table 13 depicts the path coefficients for each of the models and the characteristics of the latent constructs. The adjusted R^2 refers to the amount of variance in the outcome construct that is accounted for by the predictor constructs (i.e., fidelity dimensions) feeding into it in the model. Across outcomes, adjusted R^2 ranged from .03 to .50, with predictors (i.e., fidelity dimensions) accounting for the greatest amount of variance in resource quality and the least amount of variance in burnout and secondary trauma. The AVE assesses the validity of the factor structure for each latent construct. AVE greater than .50 is considered acceptable, reflecting that a single factor accounts for 50% or more of the variability in the items. Reliabilities are considered acceptable if greater than .70. As can be seen in Table 13, the structural fidelity, burnout, and secondary trauma measures have AVE less than .50 and lower reliabilities than do other constructs in the models; however, their reliabilities are still greater than the .70 threshold.

Table 13. Path coefficients for each of models and characteristics of latent variable composites.

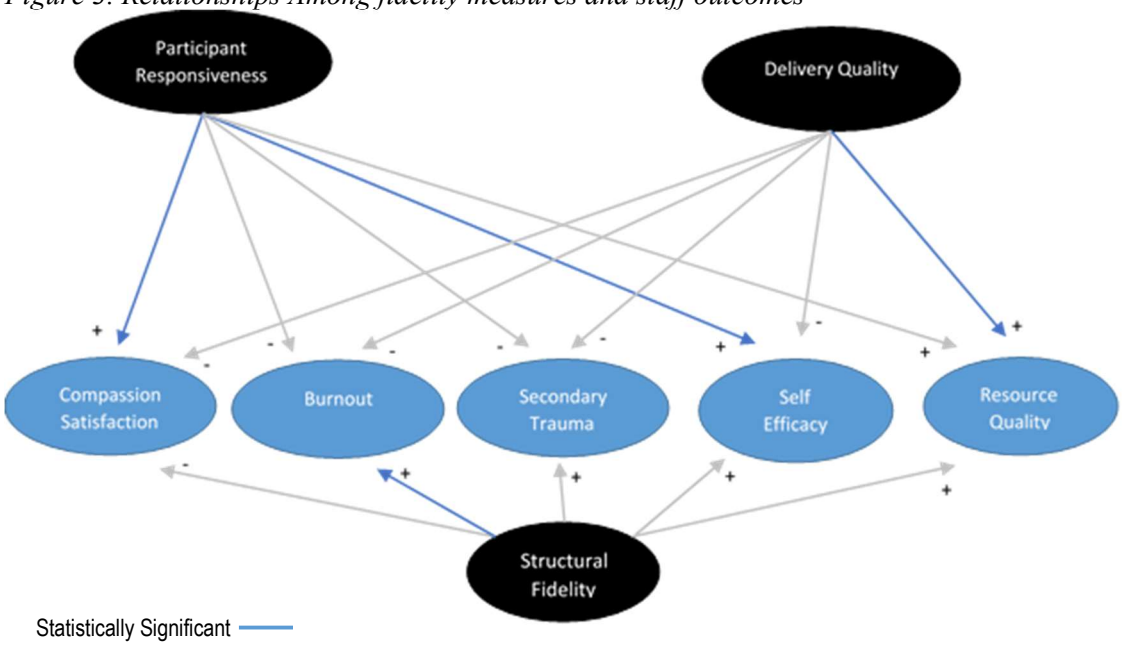
	Participant Responsiveness AVE=0.509 REL=0.940	**Delivery Quality** AVE=0.622 REL=0.961	**Structural Fidelity** AVE=0.350 REL=0.853
Resource Quality Adj R^2 = 0.50 AVE = 0.832 REL = 0.982	β =.170 se =0.10 *p* = 0.15 ES = 0.07	β = 0.53* se = 0.10 *p* < 0.002 ES = 0.36	β = 0.16 se = 0.12 p = 0.16 ES = 0.09
Self-Efficacy Adj R^2=0.18 AVE=0.476 REL=0.845	β = 0.49* se = 0.10 *p* < 0.002 ES = 0.22	β = -0.14 se = 0.12 *p* = 0.23 ES = 0.01	β = 0.05 se = 0.12 p = 0.70 ES = 0.01
Burnout Adj R^2 = 0.03 AVE = 0.267 REL = 0.763	β = -0.13 se = 0.12 *p* = 0.28 ES = 0.01	β = -0.16 se = 0.12 *p* = 0.18 ES=0.00	β = 0.32* se = 0.11 p = 0.006 ES = 0.06
Compassion Satisfaction Adj R^2 = 0.06 AVE = 0.519 REL = 0.915	β = 0.22* se = 0.11 *p* = 0.004 ES = 0.07	β = -0.18 se = 0.12 *p* = 0.12 ES = 0.02	β = -0.11 se = 0.12 p = 0.34 ES = 0.01
Secondary Trauma Adj R^2 = -0.03 AVE = 0.380 REL = 0.856	β = -0.07 se = 0.12 *p* = 0.60 ES = 0.01	β = -0.13 se = 0.12 *p* = 0.28 ES = 0.02	β = 0.05 se = 0.12 *p* = 0.66 ES = 0.00

Notes: AVE = average variance extracted. REL = composite reliability. ES = effect size, .02-.149 = small; .15-.349 = medium; .35 and higher = strong. * *p* < .05

Interpretation of Path Coefficients. The fully linked model suggests that the different aspects of fidelity make different, at least partially unique contributions to the home visitor outcomes. There was a large, significant effect between delivery quality and perceived quality of the resources. Increased participant responsiveness (i.e., home visitors' perceived efficacy actively participating in key aspects of MHC) was associated with greater self-efficacy (a medium-sized effect) and greater compassion satisfaction (a small effect) among the home visitors. Finally, an increase in adherence to the structural aspects of the fidelity model was associated with an increase in burnout (a small effect). This latter relationship suggests that home visitors may find

maintaining fidelity to the delivery model to be demanding and burdensome. However, because this is a small effect, these results suggest a need for additional exploration to understand how participation in MHC (when model adherence and exposure is greater) creates obstacles for home visitors.

Figure 3. Relationships Among fidelity measures and staff outcomes



Research Question 3

RQ3. What is the effect of Home Visitor Mental Health Consultation on parenting and family functioning/support outcomes as measured by the depression, personal care, and mobilizing resources subscales of the Healthy Families Parenting Inventory (HFPI) and the emotional and verbal responsivity of primary caregiver, organization of physical and temporal environment, and parent involvement with child subscales from the Home Observation for Measurement of the Environment Inventory (HOME)?

To address research question 3, the evaluation team employed the multistep process that is described below.

- **Step 1:** The evaluation team determined the objective (secondary activity reports) and subjective (home visitor and clinician surveys) measures of fidelity that were related to family outcomes.
- **Step 2:** The evaluation team created a summary measure of fidelity for each family drawn from measures of fidelity shown to be predictive of outcomes.
- **Step 3a:** The evaluation team examined whether the fidelity score predicts family outcomes (as measured by the HOME and HFPI scales) within the MHC treatment group.
- **Step 3b:** An exploratory quasi-experimental matched comparison groups design was conducted. MIECHV-funded families served by home visitors receiving the MHC enhancement were divided into low and high-fidelity treatment groups, and each group was compared to a separate matched group of non-MIECHV-funded families served by home visitors not receiving the MHC enhancement.

To balance treatment and comparison groups and replicate a randomized controlled trial, the evaluation team utilized propensity score matching (Rosenbaum & Rubin, 1983). While accounting for fidelity to the MHC model, this research question examined how effective MHC treatment is in improving levels of parenting and family functioning/support.

Research Question 3 – Steps 1 and 2

The goal of RQ3, Steps 1 and 2 was to identify individual fidelity measures related to improvement in family outcomes (Step 1) and to then combine those individual measures to create a single total fidelity score for each family (Step 2).

Step 1. The evaluation proposed to identify key fidelity measures related to improvement in family outcomes with multiple Mixed Linear Models (MLM), using each fidelity measure (continuous variables) to predict each of the six outcomes. In addition to this, the evaluation team also examined the ability of categorical versions of the fidelity measures to predict each of the six outcomes in separate MLMs. Furthermore, using Linear Discriminant Analyses (LDA), the evaluation team inverted the analysis to examine the ability of family outcomes to predict fidelity category (e.g., low vs. high) for that item. To simplify these analyses, the evaluation team performed preliminary MLMs to generate family-level slopes over month for each of the six outcome measures. In these preliminary MLMs, month was the sole fixed effect, and for random effects, the evaluation team modeled the slope of month for each family, nested within site. The results of this analysis produced a model-predicted slope for each family. These slopes represent the amount of improvement (or impairment) in each outcome that the family experienced over time. Then, because the HOME and HFPI subscales have different ranges, the evaluation team standardized these slope values to create variables that could be compared across the HOME and HFPI subscales. It is these standardized slopes that were used in the MLMs and LDAs performed for Step 1. Thus, the dependent measures for the Step 1 MLMs are the standardized change in outcomes over time, and for the Step 1 LDA, the predictor is the standardized change in outcome over time. A high-level summary of these results is provided here, and complete tables with the results of all analyses are provided in Appendix A.

The evaluation assessed the performance of both site-level fidelity criteria and family-level fidelity criteria. A total of nine fidelity criteria emerged as reliable predictors of at least a subset of the outcomes. Seven of these were site-level items (of a possible total of 10 site-level items) and two of these were family-level items (of a possible total of 3 family-level items). Fidelity criteria that reliably predicted outcomes appear in Table 14.

Table 14. Fidelity criteria that reliably predicted outcomes.

Fidelity Criteria	Associated Family Outcomes
Site Level Indicators	
Mean Participant Responsiveness [A]	Organization [C], Involvement [C], Responsivity [C], Mobilizing Resources [D]
Percent of families reviewed for clinical risk [B]	Organization [C], Involvement [C]
Percent of families reviewed all months [B]	Organization [C], Involvement [C], Responsivity [C]
Percent of home visitors receiving reflective practice all months [B]	Organization [C], Involvement [C], Responsivity [C]
Percent of home visitors receiving consultation all months [B]	Organization [C], Involvement [C], Responsivity [C]
Average number of total meetings per home visitor [B]	Organization [C], Involvement [C], Responsivity [C]
Total number of trainings offered [B]	Organization [C], Involvement [C], Responsivity [C]
Family Level Indicators	
Number of weeks served by MHC [B]	Personal Care [D], Mobilizing Resources [D]
Number of months reviewed by MHC [B]	Depression [D], Personal Care [D], Mobilizing Resources [D]

Data Sources: [A] Home Visitor Reflective Supervision Self-Efficacy Scale, [B] Secondary Activity Logs, [C] HOME, [D] HFPI

Step 2. With the goal of creating a single fidelity total score, the evaluation team submitted the nine fidelity items to a principal components analysis. First, exploratory parallel analysis indicated a three-factor solution would best account for the variance in the items. The evaluation team then ran a confirmatory principal components analysis with three factors and a Varimax rotation. This analysis showed that the three-component solution cumulatively explained 79% of the variance among the fidelity criteria, with the first, second and third components explaining 35%, 27%, and 16% of the variance, respectively. Typically, items are considered to be loading onto a particular component when they have a factor loading of at least 0.40. Higher factor loadings indicate greater correlation with the remainder of the items on that component. To create component scores, the evaluation team assigned the items to the component on which they loaded most heavily (see Table 15). Test-retest reliability indices for each of the components appears in Table 16. Because Cronbach's alpha can underestimate reliability, the report also presents Guttman's lambda 6, which considers the amount of variance in each item that can be accounted for via linear regression of all the other items. Both alpha and Guttman's lambda 6 values will increase with increases in the number of items and with increases in the intercorrelation among the items. Reliabilities of 0.80 or greater are preferred.

Table 15. Principal components analysis of fidelity items.

Component	Fidelity Criteria	Component Loadings		
		Component 1	Component 2	Component 3
	Site Level Indicators			
(1)	Percent of home visitors receiving reflective practice all months	.96	-.13	<.10
(1)	Percent of home visitors receiving consultation all months	.93	<.10	<.10
(1)	Average number of total meetings per home visitor	.69	.50	.19
(1)	Total number of trainings offered	.87	.30	<.10
(2)	Mean Participant Responsiveness	-.16	.80	-.15
(2)	Percent of families reviewed for clinical risk	.23	.76	.17
(2)	Percent of families reviewed all months	.23	.84	<.10
	Family Level Indicators			
(3)	Number of weeks served by MHC	<.10	-.17	.85
(3)	Number of months reviewed by MHC	<.10	-.31	.82

Table 16. Reliability for the Three Fidelity Components

Component	Component Name	Cronbach's Alpha	Guttman's Lambda 6
Component 1	Site Level Number of MHC Meetings & Trainings (4 items)	0.53	0.90
Component 2	Site Level Family Review (3 items)	0.20	0.68
Component 3	Family Level Service (2 items)	0.25	0.51

Creation of Single Total Fidelity Score per Family. The evaluation team first created family scores on each component by standardizing the family's score on a given criterion and then multiplying it by its respective component loading and then summing the resulting values for all the items on a given component. From the three component scores, the evaluation team created a single total fidelity score by multiplying each component score by the amount of variance that component explained in the overall set of items. That is, the total fidelity score = 0.35*Component1 score + 0.268*Component2 score + 0.167*Component3 score. Family fidelity scores ranged from -2.59 to 2.39 (*M* = 0.00, *SD* = 1.35). Summary statistics for family fidelity scores, summarized by county/site appear in Table 17.

Table 17. Summary statistics on family total fidelity scores, by county.

County	Mean	SD	Median	Min	Max
Site 1	1.71	0.31	1.71	1.22	2.39
Site 2	1.56	0.22	1.47	1.30	2.24
Site 3	0.30	0.32	0.23	-0.08	1.42
Site 4	-0.05	0.21	-0.09	-0.38	0.72
Site 5	-0.25	0.20	-0.30	-0.49	0.22
Site 6	-0.79	0.24	-0.88	-1.16	-0.14
Site 7	-2.26	0.18	-2.29	-2.53	-1.72
Site 8	-2.38	0.15	-2.40	-2.59	-1.95

RQ3a. Does fidelity to the MHC treatment model predict family outcomes within the MHC treatment group?

Step 3a. To determine whether fidelity to the treatment model predicted outcomes of families in the treatment group, the evaluation team performed separate MLMs on each of the three HOME and three HFPI subscales. In these models, the dependent measure was the standardized family-level slope across month on the respective subscale. Higher positive values on this measure indicate greater improvement across months, and negative values on this measure indicate a decline in the outcome across months. Using this approach, the evaluation sought to determine if increased fidelity to the model (i.e., total fidelity score) was associated with greater improvement in family outcomes across the measurement period. In essence, this is a dose-response question: if a family gets more – or a higher quality – treatment, do they have greater improvement in their outcomes?

The sample size for this analysis was *N* = 1,692 families and *N* = 8 sites. Table 18 depicts descriptive statistics for family-level slopes on each of the six subscales, as well as the intraclass correlation coefficient (ICC) for site on that subscale. The ICC indicates the amount of overall variability in the data that is accounted for by differences between the sites before the analysis begins to take account of other variables of interest. Note the greater dispersion and variability among the values for the HFPI scales (see *SD*, min, and max columns of Table 18). Inspection of the site-level intraclass correlation coefficients (ICC) suggested little variability attributable to the site-level (see Table 18). Typically, random effects are not modeled when the ICC is below 0.10. Therefore, in subsequent analyses, the evaluation team inspected both the MLM with site

as a random effect (as proposed) and the ordinary least squares (OLS) regression with no random effects.

Table 18. Descriptives on family-level slopes across month for measurement subscales and ICC for random effect of site.

Measure	Mean	SD	Min	Max	ICC
HOME subscales					
Responsivity	-0.030	0.164	-0.647	0.493	0.026
Organization	-0.014	0.078	-0.285	0.238	0.015
Involvement	-0.013	0.092	-0.361	0.262	0.029
HFPI subscales					
Depression	-0.006	1.179	-4.491	4.723	0.001
Mobilizing Resources	-0.005	0.764	-3.098	3.188	0.000
Personal Care	-0.003	0.612	-2.572	2.665	0.000

Tables 19 and 20 below depict the results of the analyses on the HOME and HFPI subscales, respectively. In the scatterplots in the bottom row of these tables, each dot represents one family's score, and the line represents how the slope of the outcome variable changed with changes in total fidelity scores. If there is no relationship between total fidelity and the outcome, these lines will be flat. If the outcomes improve with total fidelity, these lines will have a positive slope (i.e., increase from left to right). For each scale, the family-level total fidelity score resulting from steps 1 and 2 was used as the sole predictor of the standardized family-level slope across month for that subscale.

HOME Subscales. For the HOME subscales (Table 19), there were effects of total fidelity score on the slopes of the outcome measures, indicating greater improvement in the outcomes as fidelity increased (note the slight positive slope to the regression lines on the scatterplots in Table 19). These effects were statistically significant in the OLS regression for all of the HOME measures (all p-values < .001), but not always so when the random effect of site was also included in the model (see MLM results in Table 19). Regardless, in both cases, these effects were very small, with all $R^2 = 0.01$, indicating that fidelity accounted for 1% of the variance in the change in the outcome measures over time.

HFPI Subscales. For the HFPI subscales (Table 20), the results of the MLM and OLS regressions were equivalent. There were no effects of fidelity on the change in these subscales. All slopes on these measures were at zero (see flat regression lines in last row of Table 20).

Table 19. Results of analyses using total fidelity score to predict HOME subscales among MHC-treated families (Research Question 3a).

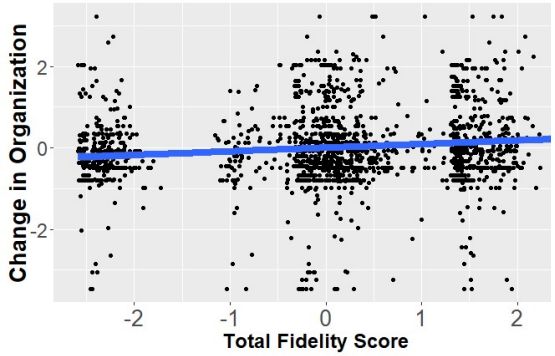
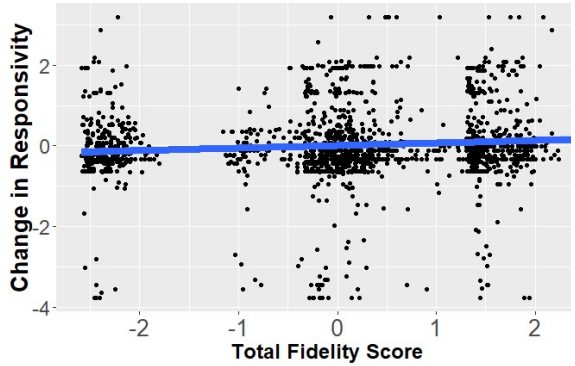
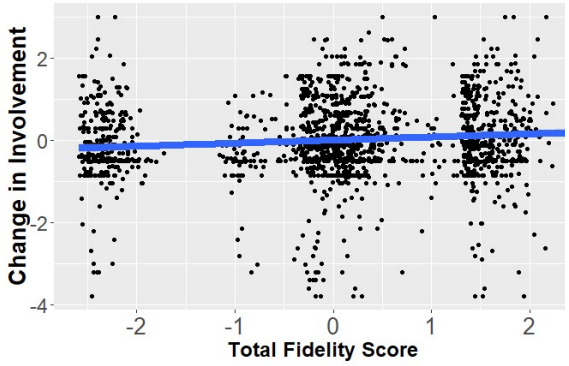
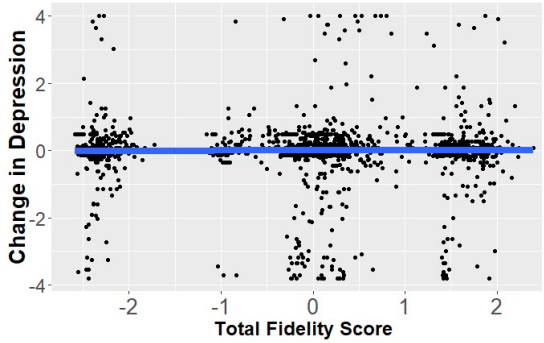
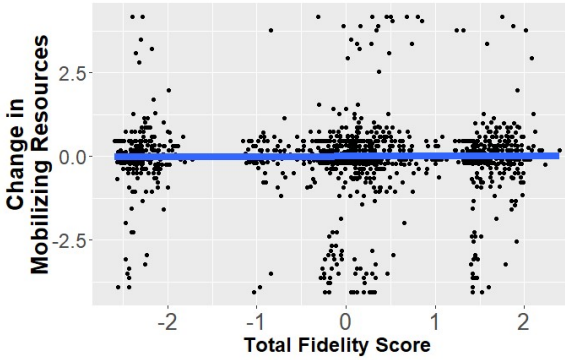
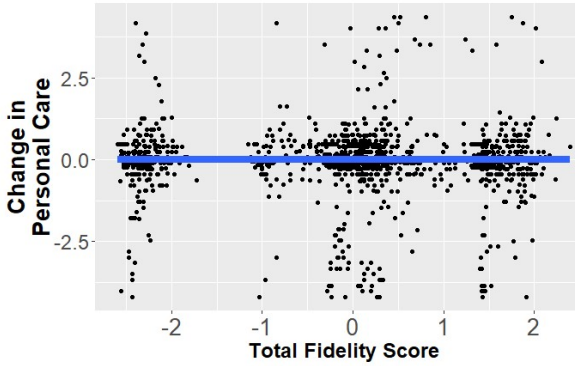
Organization Subscale	Responsivity Subscale	Involvement Subscale
MLM b = 0.06, SE = 0.03, β = 0.08, p = .066, fixed effect R^2 = 0.01, fixed + random R^2 = 0.02	***MLM*** b = 0.06, SE = 0.02, β = 0.08, p = .014, fixed effect R^2 = 0.01, fixed + random R^2 = 0.01	***MLM*** b = 0.05, SE = 0.04, β = 0.08, p = .170, fixed effect R^2 = 0.01, fixed + random R^2 = 0.02
		

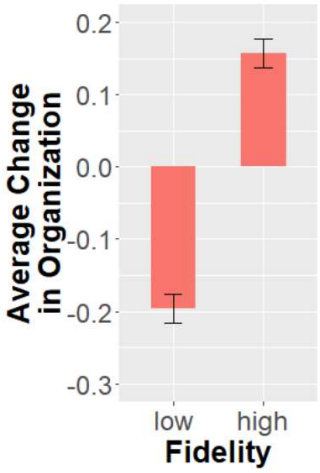
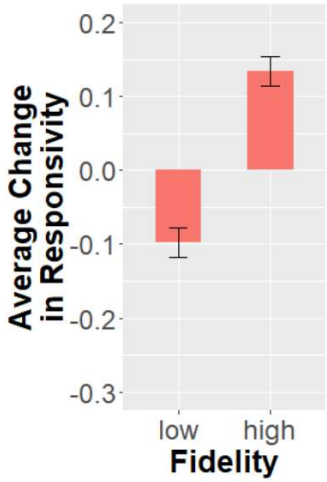
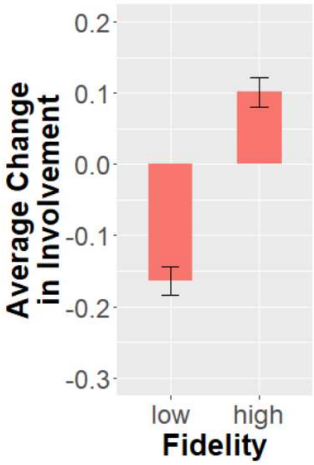
Table 20. Results of analyses using total fidelity score to predict HFPI subscales among MHC-treated families (Research Question 3a).

Depression Subscale	Mobilizing Resources Subscale	Personal Care Subscale
MLM b = 0.01, SE = 0.02, β = 0.01, p = .685, fixed effect R^2 < 0.001, fixed + random R^2 = 0.002	***MLM*** b = 0.01, SE = 0.02, β = 0.01, p = .599, fixed effect R^2 < 0.001,fixed + random R^2 < 0.001	***MLM*** b < .001, SE = 0.02, β < .001, p = .980, fixed effect R^2 < 0.001, fixed + random R^2 < 0.001
		

Exploratory Analysis of High versus Low Fidelity MHC-Treated Families. Inspection of the distribution of total fidelity score across the x-axes in the scatterplots of Tables 19 and 20 reveals three clusters of fidelity scores (those below -1, those clustering around 0, and those above 1). In an exploratory analysis, the evaluation team compared the highest and lowest fidelity families by coding those with a total fidelity score of 1 or greater as high and those with a total fidelity score of -1 or lower as low. The team then compared the mean change over time in each of six subscales across the high (n = 493) and low (n = 342) fidelity groups in separate Welch's independent t-tests. Tables 21 and 22 depict the results of these analyses for the HOME and HFPI subscales, respectively. The final row in each table depicts bar graphs of the average change in each subscale over time, with standard error bars. Negative values in these graphs indicate that the outcome got worse over time, and positive values indicate that the outcome improved over time.

HOME Subscales. The bar graphs of the HOME subscales in Table 21 clearly show that the outcomes for the lowest fidelity families actually decreased over time, while those of the highest fidelity families increased over time. These analyses confirm the results of the MLM and OLS regression above, with effect sizes here slightly larger among the HOME subscales. Larger effect sizes likely occurred because the exploratory analyses examined the differences between the *extremely* low and *extremely* high fidelity groups. The size of these effects, as assessed by Cohen's *d,* ranged from small (0.24, 0.28) to almost medium (0.37).

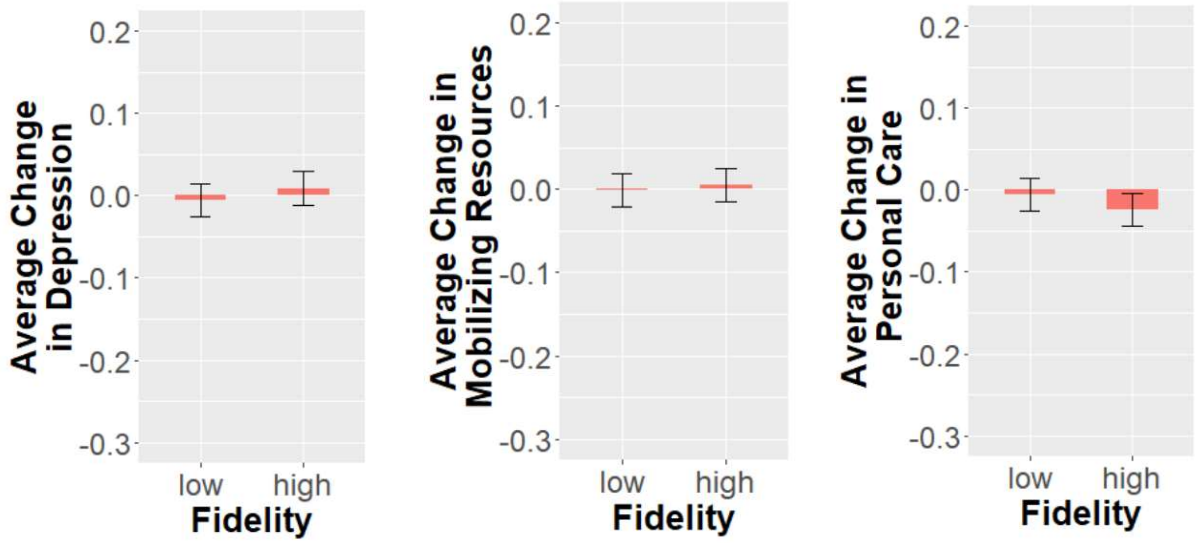
Table 21. Comparison of change in HOME outcomes in highest and lowest fidelity MHC-treated families (Research Question 3a).

Organization Subscale	Responsivity Subscale	Involvement Subscale
t = 5.47, p < .001, d = 0.37 High M = 0.16, SD = 1.03 Low M = -0.19, SD = 0.83	t = 3.62, p < .001, d = 0.24 High M = 0.13, SD = 1.03 Low M = -0.10, SD = 0.83	t = 4.03, p < .001, d = 0.28 High M = 0.10, SD = 1.01, Low M = -0.16, SD = 0.87
		

HFPI Subscales. Table 22 below depicts these same analyses for the HFPI subscales. As is apparent in the bar graphs, there was no difference in the outcomes for the low and high fidelity groups on any of the HFPI subscales. These results confirm the MLM and OLS regression analyses of the entire sample. Together, the analyses suggest that there is no association between total fidelity score and change in the HFPI outcomes over time.

Table 22. Results of analyses using total fidelity score to predict HFPI subscales among MHC-treated families (Research Question 3a).

Depression	Mobilizing Resources	Personal Care
t = 0.22, p = .824, d = 0.02 High M = 0.01, SD = 0.93 Low M = -0.01, SD = 0.89	t = 0.10, p = 0.921, d = 0.01 High M = 0.01, SD = 0.91 Low M = -0.001, SD = 0.90	t = -0.29, p = 0.771, d = -0.02 High M = -0.02, SD = 0.94 Low M = -0.01, SD = 0.90



Note on HFPI Relationships. In Step 1 of RQ3, the results showed significant associations between both family-level fidelity criteria (i.e., number of weeks served by MHC and number of months reviewed by MHC) and the HFPI subscales. Yet, in RQ3a, there was no association between the total fidelity score and the HFPI subscales. This might seem puzzling at first glance; however, when one revisits how the evaluation team created the total fidelity score, it makes sense. The total score was created by weighting the component scores and adding them together. The family-level component received the least amount of weight in creating that total score. Specifically, it was weighted at .16, whereas the other components were weighted .35 and .26. Thus, the majority of the variability in the total fidelity score is due to variability in these other components, which were related to the HOME subscales rather than the HFPI subscales.

RQ3b. Do high and low fidelity MHC treatment groups have better family outcomes than non-MHC treatment groups?

Step 3b. Evaluation Step 3b uses the propensity-matched non-treatment comparison groups that were created in Step 2 for the high and low fidelity treatment groups, based on the median split of total fidelity score. Separate sets of six 3-level MLMs (one for each of the subscales) were conducted for the high fidelity treatment group (High Fidelity MHC) and its propensity-matched non-treatment comparison group (High Fidelity Comparison), and separately for the low fidelity treatment group (Low Fidelity MHC) and its propensity-matched non-treatment comparison (Low Fidelity Comparison). In these 3-level MLMs, site was the level-3 variable, family the level-2 variable, and month the level-1 variable. Each MLM modeled the group (treatment, non-treatment) by month factorial as fixed effects and the intercepts for site and family as random effects. So, in contrast to Step 3a, here the dependent measure is the standardized score on that

outcome at that point in time (rather than standardized change in the outcome across months as in 3a).

High Fidelity Treated Families vs. Propensity-Matched Non-Treatment Comparison Families and Low Fidelity Treated Families vs. Propensity-Matched Non-Treatment Comparison Families. There were $N = 974$ families for high fidelity analyses, with $n = 487$ in the MHC treatment group and $n = 487$ in the non-treatment comparison group. In the low fidelity groups, there were a total of $N = 1558$ families, with $n = 779$ in the MHC treatment group and $n = 779$ in the non-treatment comparison group. Results for each subscale are discussed separately and then summarized. Guidelines for interpreting the tables and figures are provided in the description of the HOME subscales.

HOME Subscales. Results of the analysis of the HOME subscales appear in Table 23 and Figures 4, 5, and 6. In the table, the fixed effects represent average effects across all the observations. The slope (*b*) column for the fixed effect of group represents the overall difference between the MHC treatment and the comparison groups (averaged across all timepoints). When this value is negative, it means the comparison group was worse than the treatment group, and when this value is positive, it means the comparison group was better than the treatment group. The fixed slope (*b*) for month represents the average change in the outcome across months (averaged over group). If there were improvements in the outcomes over time, one would expect this slope to be positive, but if this slope is negative, it indicates a worsening of outcomes over time, on average. The group by month coefficient (*b*) is the important fixed effect for determining whether there is a difference in the improvement between the treated and un-treated groups. Were the MHC enhancement to influence outcomes, one would expect to see a significant group by month interaction (with the value in the *p* column $< .05$). To interpret any such interaction, one would inspect the plot of the slopes of each group in Figures 4, 5, and 6. The random effects of site and of family in Table 23 refer to variability in the outcomes attributable to differences among the sites and differences among the families. Even though there may be an average treatment effect, individual families and those at different sites will vary around the group average. The variance estimates for the random effects captures this variability.

Examining the analyses for the High Fidelity MHC group versus the High Fidelity Comparison, one sees that across the three subscales, none of the group by month interactions reach significance, which suggests no difference in the performance improvement of the two groups over the months. For the Low Fidelity MHC group versus Low Fidelity Comparison, the group by month interaction reached statistical significance for all three subscales. The pseudo-R^2 for the fixed effects components of all of the models depicted in Table 23 are less than 0.01. That is, the fixed effects in each of these models improves the ability of the model to capture the data by less than 1%.

Table 23. Results of MLM analyses for HOME subscales for MHC treatment groups and propensity matched comparison groups.

	Responsivity Subscale							
Effects	***High Fidelity MHC vs. Matched Comparison***				***Low Fidelity MHC vs. Matched Comparison***			
Fixed	***b***	***SE***	***p***	**95% *CI***	***b***	***SE***	***p***	**95% *CI***
group	-0.35	0.45	.440	-1.23, 0.53	0.06	0.49	.910	-0.90, 1.01
month	0.13	0.02	< .001	0.09, 0.18	0.03	0.02	.057	-0.00, 0.07
group X month	-0.28	0.03	.404	-0.09, 0.04	0.09	0.03	.001	0.04, 0.14
Random	**Var Est**	***SE***		**95% *CI***	**Var Est**	***SE***		**95% *CI***
site	0.79	0.40		0.29, 2.23	0.88	0.31		0.44, 1.74
family	2.90	0.36		2.27, 3.69	3.42	0.29		0.04, 0.14

	Organization Subscale							
Effects	***High Fidelity MHC vs. Matched Comparison***				***Low Fidelity MHC vs. Matched Comparison***			
Fixed	***b***	***SE***	***p***	**95% *CI***	***b***	***SE***	***p***	**95% *CI***
group	-0.73	0.26	.006	-1.25, -0.21	0.10	0.24	.678	-0.38, 0.58
month	0.07	0.12	<.001	0.047, 0.09	0.02	0.01	.019	0.00, 0.04
group X month	-0.012	0.02	.483	-0.05, 0.02	0.04	0.01	.002	0.02, 0.07
Random	**Var Est**	***SE***		**95% *CI***	**Var Est**	***SE***		**95% *CI***
site	0.33	0.18		0.11, 0.96	0.21	0.07		0.10, 0.42
family	0.92	0.11		0.72, 1.16	1.06	0.09		0.90, 1.25

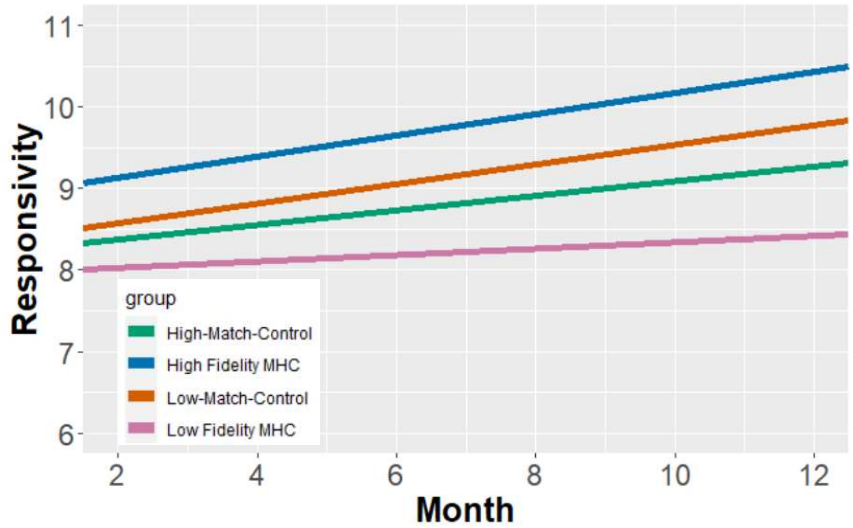
	Involvement Subscale							
Effects	***High Fidelity MHC vs. Matched Comparison***				***Low Fidelity MHC vs. Matched Comparison***			
Fixed	***b***	***SE***	***p***	**95% *CI***	***b***	***SE***	***p***	**95% *CI***
group	-0.14	0.26	.582	-0.66, 0.37	0.10	0.26	.699	-0.41, 0.61
month	0.09	0.01	<.001	0.07, 0.12	0.05	0.01	<.001	0.03, 0.07
group X month	-0.02	0.02	.408	-0.06, 0.02	0.04	0.01	.008	0.01, 0.07
Random	**Var Est**	***SE***		**95% *CI***	**Var Est**	***SE***		**95% *CI***
site	0.26	0.14		0.09, 0.73	0.24	0.09		1.16, 1.55
family	1.17	0.12		0.95, 1.43	1.34	0.10		1.16, 1.55

One can inspect Figures 4 through 6 to better understand how the HOME outcomes change over time in each of the groups. In each of the Figures, the statistics in the right-hand column are the slopes (*b*) describing change over time for each group separately. Positive slopes indicate improvement, and negative slopes a worsening in the outcome. *p*-values less than 0.05 indicate a slope that is significantly different from zero. In each of these figures, the raw averages for each group at each month are depicted in a summary table. Sometimes the trend visible in the raw data differs slightly from that in the MLM predicted slopes. This occurs because the MLM model takes into account the random variability among the sites and families, and thus, the MLM slopes represent only the fixed effect of group by month. The raw data, however, do not separate out the random effects of site and family.

HOME – Responsivity. As shown in Figure 4, all groups showed significant improvement in responsivity over time, with the exception of the Low Fidelity MHC group. The High Fidelity MHC group (blue line) and the High Fidelity Comparison group (green line) had very similar slopes of 0.13 and 0.11, respectively. The lack of a significant group by month interaction for responsivity in Table 23 indicates this small slope difference was not significant. The Low Fidelity Comparison group (orange line) also performed similarly, with a slope of 0.12. The Low Fidelity MHC group (pink line), however, had a slope of only 0.03, which did differ significant from zero. Given, the significant month by group interaction shown in Table 23, the Low Fidelity MHC group showed less improvement compared to the Low Fidelity Comparison group.

Figure 4. Responsivity subscale – MLM Slope Parameters and Raw Data Descriptives

MLM Slope Parameters



High Fidelity Comparison
b = 0.11, *se* = .03, *p* < .001, 95% *CI* [0.06, 0.16]

High Fidelity MHC
b = 0.13, *se* = .02, *p* < .001, 95% *CI* [0.09, 0.18]

Low Fidelity Comparison
b = 0.12, *se* = .02, *p* < .001, 95% *CI* [0.09, 0.16]

Low Fidelity MHC
b = 0.03, *se* = .02, *p* = .057, 95% *CI* [-.001, 0.07]

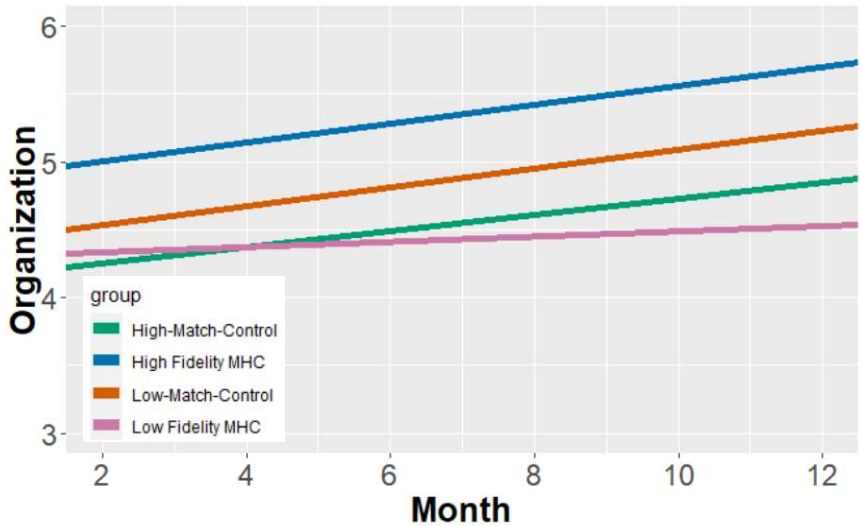
Summary Statistics of the Raw Data for Each Group by Month

RESPONSIVITY					
		MHC Treated		Matched Comparison	
Group	**Month**	**N**	**Mean (SD)**	**N**	**Mean (SD)**
High Fidelity	2	441	8.05 (3.83)	435	7.92 (3.73)
High Fidelity	4	462	9.16 (3.03)	415	8.39 (3.68)
High Fidelity	12	298	9.79 (2.7)	222	9.43 (3.32)
Low Fidelity	2	698	8.44 (3.43)	709	8.32 (3.53)
Low Fidelity	4	740	8.94 (3.32)	679	8.66 (3.56)
Low Fidelity	12	393	9.27 (3.44)	344	9.92 (2.58)

HOME – Organization. The results for the organization subscale (Figure 5) of the HOME are very similar to those for responsivity. As with responsivity, the High Fidelity MHC group, the High Fidelity Comparison group, and the Low Fidelity Comparison group all performed similarly, showing slopes in the 0.06 to 0.07 range that were significantly different from zero. The Low Fidelity MHC group, however, while also having a slope that was significantly greater than zero, demonstrated much less improvement over time than observed in the other groups.

Figure 5. Organization subscale – MLM Slope Parameters and Raw Data Descriptives

MLM Slope Parameters



High Fidelity Comparison
b = 0.06, *se* = .01, *p* < .001, 95% *CI* [0.03, 0.08]

High Fidelity MHC
b = 0.07, *se* = .01, *p* < .001, 95% *CI* [0.05, 0.09]

Low Fidelity Comparison
b = 0.07, *se* = .01, *p* < .001, 95% *CI* [0.04, 0.09]

Low Fidelity MHC
b = 0.02, *se* = .01, *p* = .019, 95% *CI* [-.003, 0.04]

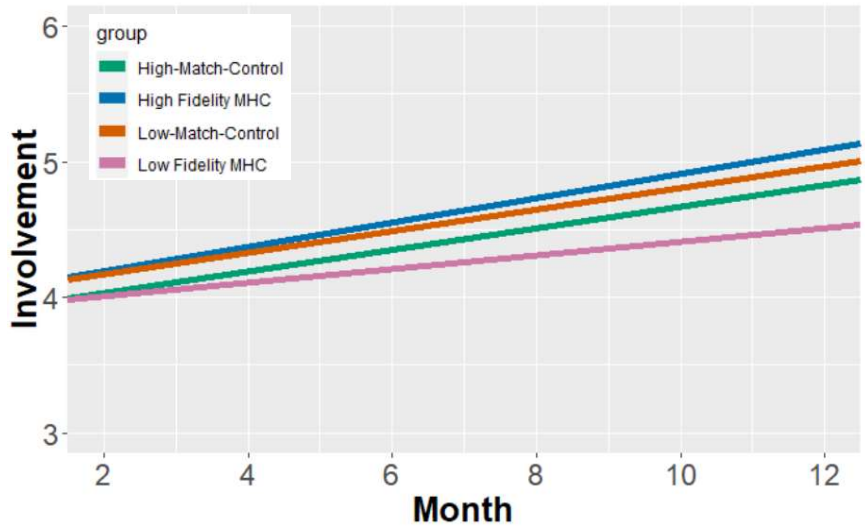
Summary Statistics of the Raw Data for Each Group by Month

ORGANIZATION					
		MHC Treated		Matched Comparison	
Group	**Month**	**N**	**Mean (SD)**	**N**	**Mean (SD)**
High Fidelity	2	441	4.15 (2.07)	435	4.17 (2.08)
High Fidelity	4	462	4.75 (1.73)	415	4.51 (2.05)
High Fidelity	12	298	5.06 (1.58)	222	5.00 (1.80)
Low Fidelity	2	698	4.43 (1.89)	709	4.34 (1.97)
Low Fidelity	4	740	4.66 (1.84)	679	4.56 (1.99)
Low Fidelity	12	393	4.90 (1.9)	344	5.19 (1.47)

HOME – Involvement. For the involvement subscale of the HOME (depicted below in Figure 6), all groups showed significant improvement over time, but again, the Low Fidelity MHC group showed the least amount of improvement, having the smallest slope of the four groups (*b* = 0.05).

Figure 6. Involvement subscale – MLM Slope Parameters and Raw Data Descriptives

MLM Slope Parameters



High Fidelity Comparison
b = 0.08, *se* = .01, *p* < .001, 95% *CI* [0.05, 0.10]

High Fidelity MHC
b = 0.09, *se* = .01, *p* < .001, 95% *CI* [0.07, 0.12]

Low Fidelity Comparison
b = 0.08, *se* = .01, *p* < .001, 95% *CI* [0.06, 0.10]

Low Fidelity MHC
b = 0.05, *se* = .01, *p* < .001, 95% *CI* [0.03, 0.07]

Summary Statistics of the Raw Data for Each Group by Month

INVOLVEMENT					
		MHC Treated		Matched Comparison	
Group	**Month**	**N**	**Mean (SD)**	**N**	**Mean (SD)**
High Fidelity	2	441	3.63 (2.09)	435	3.82 (2.19)
High Fidelity	4	462	4.41 (1.78)	415	4.28 (2.12)
High Fidelity	12	298	4.79 (1.68)	222	4.86 (1.87)
Low Fidelity	2	698	3.96 (2.01)	709	4.05 (2.07)
Low Fidelity	4	740	4.35 (1.95)	679	4.37 (2.06)
Low Fidelity	12	393	4.70 (1.93)	344	5.09 (1.54)

HFPI Subscales. Results of the analysis of the HFPI subscales appear in Table 24 and Figures 7, 8, and 9. The results for the HFPI subscales are similar to those for the HOME subscales. There were no significant interactions between group and month for the High Fidelity MHC group and the High Fidelity Comparison group, but there were significant interactions between group and month for the Low Fidelity MHC group and the Low Fidelity Comparison group. As with the HOME subscales, this interaction in the low fidelity analysis reflects the Low Fidelity MHC group showing poorer improvement in outcomes across the months relative to the Low Fidelity Comparison group (see Figures 7 through 9).

Table 24. Results of analyses for HFPI subscales for MHC treatment groups and propensity matched comparison groups.

	Depression Subscale							
Effects	***High Fidelity MHC vs. Matched Comparison***				***Low Fidelity MHC vs. Matched Comparison***			
Fixed	***b***	***SE***	***p***	**95% *CI***	***b***	***SE***	***p***	**95% *CI***
group	-0.41	1.61	.799	-3.56, 2.74	1.99	2.02	.326	-1.98, 5.95
month	-0.00	0.09	.982	-0.19, 0.19	-0.34	0.08	<.001	-0.50, -0.19
group X month	-0.05	0.15	.751	-0.33, 0.24	0.28	0.12	.017	0.05, 0.51
Random	**Var Est**	***SE***		**95% *CI***	**Var Est**	***SE***		**95% *CI***
site	4.64	3.09		1.25, 17.12	14.00	4.82		7.14, 27.49
family	58.67	6.41		47.36, 72.68	71.61	5.46		61.67, 83.16

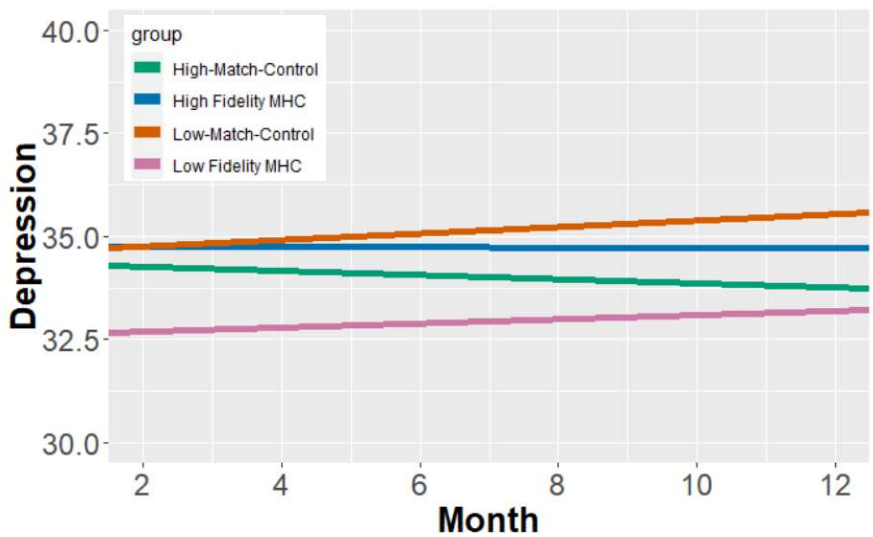
	Mobilizing Resources Subscale							
Effects	***High Fidelity MHC vs. Matched Comparison***				***Low Fidelity MHC vs. Matched Comparison***			
Fixed	***b***	***SE***	***p***	**95% *CI***	***b***	***SE***	***p***	**95% *CI***
group	0.21	1.10	.850	-1.96, 2.38	0.81	1.34	.547	-1.82, 3.43
month	0.11	0.06	.072	-0.01, 0.23	-0.17	0.05	.001	-0.27, -0.06
group X month	-0.08	0.09	.397	-0.26, 0.10	0.25	0.08	.001	0.10, 0.40
Random	**Var Est**	***SE***		**95% *CI***	**Var Est**	***SE***		**95% *CI***
site	2.61	1.64		0.76, 8.94	6.22	2.06		3.25, 11.92
family	29.44	2.81		24.42, 35.49	30.54	2.31		26.33, 35.42

	Personal Care Subscale							
Effects	***High Fidelity MHC vs. Matched Comparison***				***Low Fidelity MHC vs. Matched Comparison***			
Fixed	***b***	***SE***	***p***	**95% *CI***	***b***	***SE***	***p***	**95% *CI***
group	-0.42	0.87	.629	-2.13, 1.29	0.70	0.98	.449	-1.18, 2.66
month	0.02	0.05	.720	-0.08, 0.12	-0.16	0.04	<.001	-0.24, -0.07
group X month	-0.04	0.08	.551	-0.19, 0.10	0.16	0.06	.009	0.04, 0.28
Random	**Var Est**	***SE***		**95% *CI***	**Var Est**	***SE***		**95% *CI***
site	1.49	0.99		0.40, 5.47	3.05	1.09		1.51, 6.19
family	18.42	1.79		15.22, 22.30	21.92	1.54		19.10, 25.17

HFPI – Depression. For the depression subscale of the HFPI, however, there was a slightly different pattern than observed for the HOME subscales when the slopes for the individual groups are examined. In this case, the slopes of the High Fidelity MHC group and High Fidelity Comparison group were both negative, although only significantly so for the High Fidelity Comparison group. This negative slope suggests that depression for the High Fidelity Comparison group worsened over time (see green line in graph). Both the Low Fidelity Comparison group and the Low Fidelity MHC group had positive slopes, demonstrating improvement in depression over time, with the slope of the Low Fidelity Comparison group (orange line) slightly better than that of the Low Fidelity MHC group (pink line). It is unclear why there was a paradoxical change in depression scores over time for the High Fidelity MHC group and the High Fidelity Comparison group. Before conclusions may be drawn about this relationship, additional evaluation is recommended (see *XII.G Recommendations*).

Figure 7. Depression subscale – MLM Slope Parameters and Raw Data Descriptives

MLM Slope Parameters



High Fidelity Comparison
b = -0.05, *se* = .01, *p* < .001, 95% *CI* [0.03, 0.08]

High Fidelity MHC
b = -0.002, *se* = .01, *p* = .982 95% *CI* [-0.19, 0.19]

Low Fidelity Comparison
b = 0.08, *se* = .01, *p* < .001, 95% *CI* [0.06, 0.10]

Low Fidelity MHC
b = 0.05, *se* = .01, *p* < .001, 95% *CI* [0.03, 0.07]

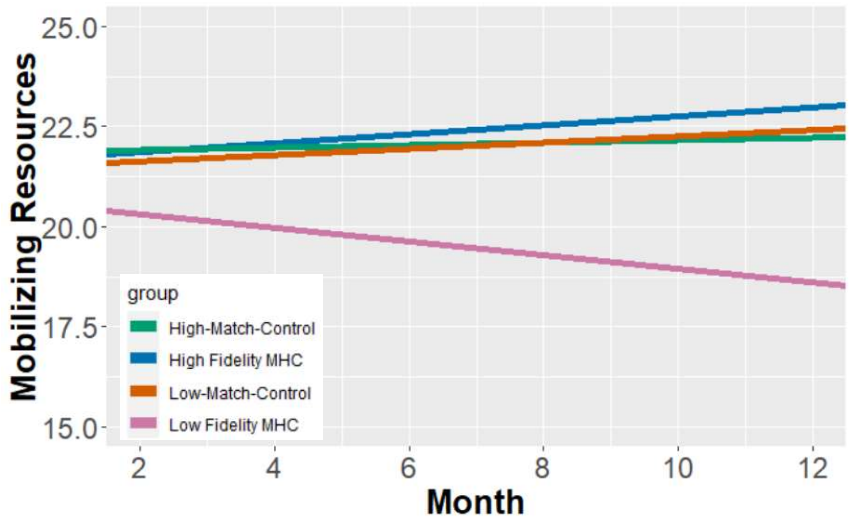
Summary Statistics of the Raw Data for Each Group by Month

DEPRESSION					
		MHC Treated		Matched Comparison	
Group	**Month**	**N**	**Mean (SD)**	**N**	**Mean (SD)**
High Fidelity	3	460	33.33 (13.68)	457	33.46 (13.56)
High Fidelity	6	473	32.43 (15.17)	340	34.24 (13.34)
High Fidelity	12	309	34.60 (12.61)	218	34.63 (13.09)
Low Fidelity	3	726	33.83 (13.08)	733	33.97 (13.47)
Low Fidelity	6	756	30.30 (16.38)	576	33.64 (14.13)
Low Fidelity	12	397	33.73 (14.18)	355	35.26 (12.25)

HFPI – Mobilizing Resources. Figure 8 depicts the change over time on the mobilizing resources subscale of the HFPI. Here again, the data show the Low Fidelity MHC group performing worse than the other three groups. In this case, the slopes of the other three groups were positive, but none were significantly different from zero (all *p*s > .05). However, the slope of the Low Fidelity MHC group was negative, suggesting a decrease in the outcome over time. It should be noted that the slope depicted in Figure 8 differs from the raw data in the table beneath it due to the MLM model accounting for the random effects of family and site.

Figure 8. Mobilizing Resources subscale – MLM parameters and raw descriptives.

MLM Slope Parameters



High Fidelity Comparison
b = 0.03, *se* = .07, *p* = .661,
95% *CI* [-0.11, 0.17]

High Fidelity MHC
b = 0.11, *se* = .06, *p* = .072
95% *CI* [-0.01, 0.23]

Low Fidelity Comparison
b = 0.08, *se* = .06, *p* = .134,
95% *CI* [-0.03, 0.19]

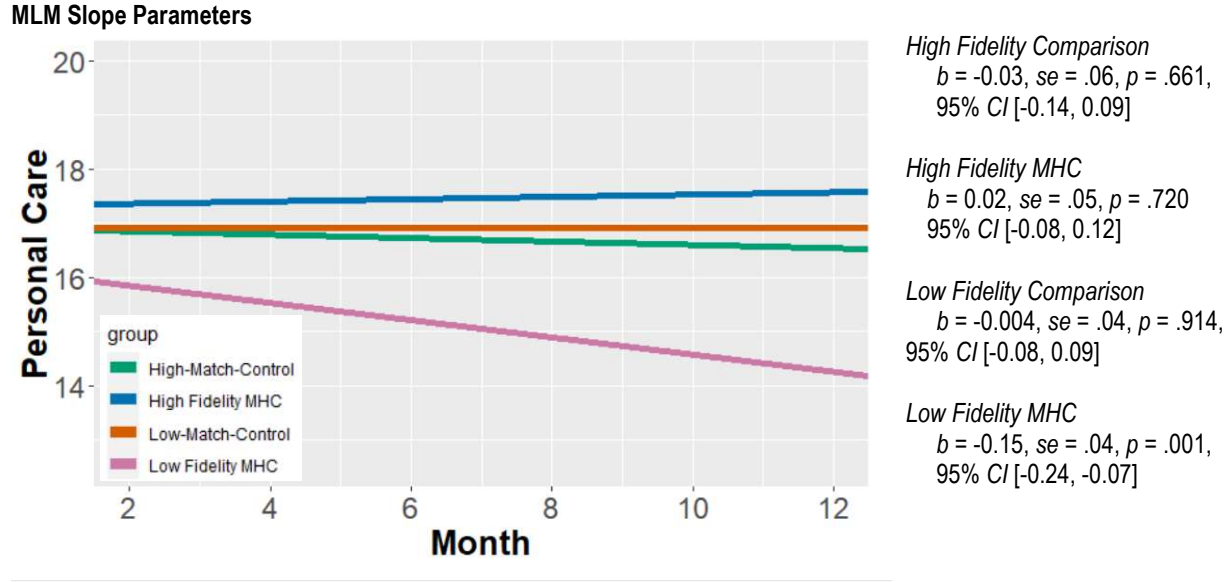
Low Fidelity MHC
b = -0.17, *se* = .05, *p* = .001,
95% *CI* [-0.27, -0.06]

Summary Statistics of the Raw Data for Each Group by Month

MOBILIZING RESOURCES					
		MHC Treated		Matched Comparison	
Group	**Month**	**N**	**Mean (SD)**	**N**	**Mean (SD)**
High Fidelity	3	460	20.56 (8.93)	457	21.17 (9.03)
High Fidelity	6	473	20.53 (9.92)	340	22.48 (8.83)
High Fidelity	12	309	22.50 (8.49)	218	22.66 (8.75)
Low Fidelity	3	726	21.50 (8.65)	733	21.35 (8.85)
Low Fidelity	6	756	19.48 (10.77)	576	21.80 (9.31)
Low Fidelity	12	397	21.91 (9.16)	355	23.28 (8.15)

HFPI – Personal Care. The personal care subscale of the HFPI is depicted in Figure 6. The results for this subscale are similar to those for the mobilizing resources subscale. The Low Fidelity MHC group had a significant, negative slope (pink line), suggesting a worsening of the outcome over time, while all of the other groups had flat slopes that did not differ significantly from zero.

Figure 9. Personal Care subscale – MLM slope parameters and raw descriptives



Summary Statistics of the Raw Data for Each Group by Month

PERSONAL CARE					
		MHC Treated		Matched Comparison	
Group	**Month**	**N**	**Mean (SD)**	**N**	**Mean (SD)**
High Fidelity	3	460	16.57 (7.3)	457	16.4 (7.15)
High Fidelity	6	473	16.11 (7.98)	340	16.81 (7.1)
High Fidelity	12	309	17.44 (6.93)	218	17 (7.17)
Low Fidelity	3	726	16.8 (7.13)	733	16.71 (7.15)
Low Fidelity	6	756	15.11 (8.54)	576	16.79 (7.52)
Low Fidelity	12	397	16.85 (7.42)	355	17.7 (6.74)

B. Sample and Comparison Group

Research Question 1

Mental Health Consultant Interviews. Interview responses were obtained from seven of the mental health consultants who provided MHC at HFI sites serving MIECHV-funded families during the fall 2019. Participants represented all seven sites that were served by mental health consultants during the evaluation window (with two mental health consultants serving Lake County and the remaining five mental health consultants serving one site each). Experience ranged from less than one year to three to five years: less than one year (29%, *n* = 2), one to two years (57%, *n* = 4), and three to five years (14%, *n* = 1).

Program Manager Interviews. Interview responses were obtained from all seven of the program managers who worked at HFI sites serving MIECHV-funded families (using the MIECHV MHC model) during the fall 2019. Participants represented all seven sites that were served by mental health consultants during the evaluation window. Experience ranged from less than one year to more than ten years: less than one year (17%, *n* = 1), one to two years (50%, *n* = 3), five to ten years (33%, *n* = 2), and more than ten years (17%, *n* = 1).

Research Question 2

The sample was drawn from the full population of home visitors who receive MHC at HFI sites serving MIECHV-funded families (using the MIECHV MHC model). The evaluation team provided a survey link to the DCS designee who forwarded the link to HFI Site Managers for administration to the home visitors. Participants received a link to electronic surveys via email. Survey links were sent to 92 home visitors receiving MHC, and 74 completed at least one scale item (response rate = 80.4%).

Participant county/program of employment was proportional to the total number of home visitors per site: Site 1 (4%; *n* = 3), Site 2 (12%; *n* = 9), Site 3 (13%; *n* = 10), Site 4 (16%; *n* = 12), Site 5 (3%; *n* = 2), Site 6 (9%; *n* = 7), Site 7 (22%, *n* = 17), and Site 8 (20%; *n* = 15). Two respondents did not indicate the site in which they worked.

Table 25. Home visitor demographics.

Demographics		
Characteristic	***n***	**%**
Years of Home Visiting Experience		
Less than 1 year	17	23%
1-2 years	20	27%
3-5 years	16	21%
5-10 years	11	15%
More than 10 years	11	15%
Sex		
Female	75	99%
Male	1	1%
Race		
American Indian or Native Alaskan	1	1%
Asian	1	1%
Black or African American	13	18%
Multiracial	4	6%
White	54	74%
Other	1	1%
Ethnicity		

Characteristic	Demographics	n	%
	Non-Hispanic	60	81%
	Hispanic	14	19%
Age			
	20-29	30	39%
	30-39	16	21%
	40-49	12	16%
	50-59	17	22%
	60 or older	2	3%
Education			
	Some high school, no degree		
	High school/GED	2	3%
	Vocational/technical training program	2	3%
	Some college, no degree	6	8%
	Associate's degree	5	7%
	Bachelor's degree	56	75%
	Master's degree	3	4%
	Professional degree	1	1%
Field of Study[A]			
	Child Development	10	13%
	Early Childhood Education	10	13%
	Education	7	9%
	Nursing	2	3%
	Psychology	19	25%
	Social Work/Social Welfare	28	37%
	Other	27	36%

Note: [A] Select all that apply.

Research Question 3

Samples utilized for RQ3 were drawn from the full population of families participating in HFI from January 1, 2016 to December 31, 2019 for whom necessary outcome data were available. This consisted of 1,692 MIECHV-funded families receiving the MHC enhancement and 6,251 non-MIECHV-funded families.

Summary of MHC cases. The MLMs proposed for RQ3 required at least one datapoint for each subscale across the three months assessed. As a result, the sample size for most analyses involving just the MHC-treated group was $N = 1{,}692$. Table 26 provides the total N, complete cases, and proportion of missing values per family for each of the counties.

Table 26. Number of Families per County/Site

County/Site	Total *N*	Complete Cases	Proportion of Missing Values Per Family			
			Mean	Median	Min	Max
Site 1	78	36	0.22	0.23	0.00	0.67
Site 2	985	381	0.26	0.33	0.00	0.87
Site 3	865	362	0.24	0.33	0.00	1.00
Site 4	626	245	0.25	0.33	0.00	1.00
Site 5	50	20	0.25	0.30	0.00	0.67
Site 6	203	111	0.20	0.00	0.00	0.87
Site 7	548	326	0.17	0.00	0.00	1.00
Site 8	370	158	0.25	0.27	0.00	1.00

RQ3a - Higher and Lower Fidelity Intervention Families. After the fidelity measures predicting family outcomes were identified and fidelity scores created, the evaluation team separated families receiving the enhancement into a higher or lower fidelity group using a median split. From the full sample of 1,692 MIECHV-funded families receiving the enhancement, 846 lower fidelity families and 846 higher fidelity families were identified.

RQ3b - Matched Groups - Low Fidelity MHC Families. A total sample of 1,558 families (779 lower fidelity intervention families and 779 comparison families) were identified through propensity score matching, and the matching process yielded balanced samples based on the covariates of interest. Multivariate and univariate tests revealed no evidence of imbalance. The overall balance chi-square test (Hansen & Bowers, 2010) was nonsignificant, which indicated that no variable or linear combination of variables was significantly unbalanced after matching. Additionally, there were no statistically significant differences between the intervention and comparison group on any individual covariates after matching. No standardized differences between treatment and comparison means exceeded .13 for any covariates, which indicated small differences between groups following matching and was consistent with recent recommendations (Ho, Imai, King, & Stuart, 2007). Of note, balance was achieved on the covariates required for a moderate study rating (the highest possible for a matched comparison group design) from HomVee (2014): 1) race/ethnicity ($p = .93$, $d < .04$ for each category), 2) socioeconomic status (income: $p = .65$, $d = .03$), and 3) baseline outcomes (HFPI: $p = .73$, $d = .02$; HOME: $p = .48$, $d = .04$). Demographic characteristics for the intervention and comparison groups are presented below.

Table 27. Demographics for low fidelity MHC families and their matched comparison group.

Demographics Characteristic	Intervention Group (*N* = 779)		Comparison Group (*N* = 779)	
Race/Ethnicity[1]	*n*	%	*n*	%
African	10	1%	10	1%
African American not of Hispanic Origin	252	32%	241	31%
American Indian/Alaskan Native	0	0%	0	0%
Anglo/American	330	42%	342	44%
Asian/Pacific Islander	15	2%	16	2%
English	1	0%	1	0%
Hispanic/Latino	120	15%	110	14%
Multi-ethnic	45	6%	48	6%
Spanish	0	0%	1	0%
Other	6	1%	10	1%
Income[1]	*n*	*M (SD)*	*n*	*M (SD)*
Annual Household Income	760	$10,784.14 ($13,041.39)	759	$10,501.77 ($11,815.49)
Caregiver Age	*n*	*M (SD)*	*n*	*M (SD)*
Age in Years at Consent Date	774	26.10 (5.81)	770	25.64 (6.32)
Education	*n*	%	*n*	%
Less than HS	79	10%	81	10%
High School Diploma or GED	99	13%	87	11%
Some College	88	11%	94	12%
Four Year Degree or Higher	21	3%	22	3%
Unknown	492	63%	495	64%
Language Spoken in the Home	*n*	%	*n*	%
English	714	92%	711	91%
Spanish	49	6%	47	6%
Other	16	2	21	3%
County Type (USDA Urban Influence Code)	*n*	%	*n*	%
Large-in a metro area with at least 1 million residents or more	339	44%	357	46%

Demographics Characteristic		Intervention Group (*N* = 779)		Comparison Group (*N* = 779)	
Small-in a metro area with fewer than 1 million residents		419	54%	387	50%
Micropolitan adjacent to a large metro area		21	3%	12	2%
Noncore adjacent to a large metro area		0	0%	5	1%
Micropolitan adjacent to a small metro area		0	0%	11	1%
Noncore adjacent to a small metro with town of at least 2,500 residents		0	0%	3	0%
Micropolitan not adjacent to a metro area		0	0%	3	0%
Noncore adjacent to micro area and contains a town of 2,500-19,999 residents		0	0%	1	0%
History of Mental Illness		*n*	%	*n*	%
	No	504	65%	502	64%
	Yes	275	35%	277	36%
History of Criminality		*n*	%	*n*	%
	No	665	85%	664	85%
	Yes	114	15%	115	15%
Alcohol Use Pre-Pregnancy		*n*	%	*n*	%
	No	748	96%	746	96%
	Yes	29	4%	31	4%
Limited Use of Alcohol by Mother		*n*	%	*n*	%
	No	136	19%	143	20%
	Yes	598	82%	592	81%
Alcohol Use Post-Pregnancy		*n*	%	*n*	%
	No	755	97%	759	98%
	Yes	21	3%	17	2%
Drug Use During Pregnancy		*n*	%	*n*	%
	No	668	86%	670	87%
	Yes	108	14%	104	13%
Drug Use Before Pregnancy		*n*	%	*n*	%
	No	192	26%	201	27%
	Yes	543	74%	536	73%
Birth Status of Target Child		*n*	%	*n*	%
	Not First Birth	447	57%	428	55%
	First Birth	332	43%	351	45%
Intimate Partner Violence Status		*n*	%	*n*	%
	No Intimate Partner Violence	481	63%	479	63%
	Intimate Partner Violence	278	37%	285	37%
Edinburgh Postnatal Depression Scale		*n*	*M (SD)*	*n*	*M (SD)*
	Initial EPDS Scores	768	5.77 (5.05)	767	5.78 (4.836)
Baseline HFPI Score[1]		*n*	*M (SD)*	*n*	*M (SD)*
	3-Month Total HFPI Score	724	234.78 (87.92)	731	234.44 (90.92)
Baseline HOME Score[1]		*n*	*M (SD)*	*n*	*M (SD)*
	2-Month Total HOME Score	702	30.44 (12.71)	711	30.64 (13.11)

[1] Note: The baseline equivalence requirement for HomVEE's (2014) moderate study rating (i.e., non-significant differences between groups on race/ethnicity, socioeconomic status, and baseline measures of outcomes) was met.

RQ3b - Matched Groups - High Fidelity Families. A total sample of 974 families (487 high fidelity intervention families and 487 comparison families) were identified through propensity score matching, and the matching process yielded balanced samples based on the covariates of interest. Multivariate and univariate tests revealed no evidence of imbalance. The overall balance chi-square test (Hansen & Bowers, 2010) was nonsignificant, which indicated that no variable or linear combination of variables was significantly unbalanced after matching. Additionally, there were no statistically significant differences between the intervention and comparison group on any covariates after matching. No standardized differences between treatment and comparison group means exceeded .16 for any covariates, which indicated small differences between groups

following matching and was consistent with recent recommendations (Ho, Imai, King, & Stuart, 2007). Of note, balance was achieved on the covariates required for a moderate study rating (the highest possible for a matched comparison group design) from HomVee (2014): 1) race/ethnicity (p = .89, d < .07 for each category), 2) socioeconomic status (income: p = .64, d = .03), and 3) baseline outcomes (HFPI: p = .76, d = .02; HOME: p = .72, d = .02). Demographic characteristics for the intervention and comparison groups are presented below.

Table 28. Demographics for high fidelity MHC families and their matched comparison group.

Demographics				
Characteristic	**Intervention Group (*N* = 487)**		**Comparison Group (*N* = 487)**	
Race/Ethnicity[1]	*n*	%	*n*	%
African	12	3%	9	2%
African American not of Hispanic Origin	135	28%	120	25%
American Indian/Alaskan Native	2	0%	3	1%
Anglo/American	218	45%	23	47%
Asian/Pacific Islander	7	1%	9	2%
Hispanic/Latino	91	19%	88	18%
Multi-ethnic	15	3%	19	4%
Spanish	0	0%	1	0%
Other	7	1%	8	2%
Income[1]	*n*	*M (SD)*	*n*	*M (SD)*
Annual Household Income	476	$11,482.72 ($12,707.21)	472	$11,960.73 ($12,192.70)
Caregiver Age	*n*	*M (SD)*	*n*	*M (SD)*
Age in Years at Consent Date	487	27.08 (6.33)	483	26.54 (6.05)
Education	*n*	%	*n*	%
Less than HS	42	9%	49	10%
High School Diploma or GED	61	13%	59	12%
Some College	49	10%	48	10%
Four Year Degree or Higher	8	2%	8	2%
Unknown	327	67%	323	66%
Language Spoken in the Home	*n*	%	*n*	%
English	421	86%	424	87%
Spanish	56	12%	53	11%
Other	10	2%	10	2%
County Type (USDA Urban Influence Code)	*n*	%	*n*	%
Large-in a metro area with at least 1 million residents or more	484	99%	476	98%
Small-in a metro area with fewer than 1 million residents	0	0%	3	1%
Micropolitan adjacent to a large metro area	3	1%	1	0%
Noncore adjacent to a large metro area	0	0%	1	0%
Micropolitan adjacent to a small metro area	0	0%	3	1%
Noncore adjacent to a small metro with town of at least 2,500 residents	0	0%	3	1%
History of Mental Illness	*n*	%	*n*	%
No	321	66%	322	66%
Yes	166	34%	165	34%
History of Criminality	*n*	%	*n*	%
No	422	87%	423	87%
Yes	65	13%	64	13%
Alcohol Use Pre-Pregnancy	*n*	%	*n*	%
No	464	96%	468	97%
Yes	20	4%	16	3%
Limited Use of Alcohol by Mother	*n*	%	*n*	%
No	118	25%	113	25%
Yes	334	75%	345	75%
Alcohol Use Post-Pregnancy	*n*	%	*n*	%
No	469	97%	468	97%
Yes	14	3%	16	3%
Drug Use During Pregnancy	*n*	%	*n*	%
No	433	90%	434	90%

Characteristic	Demographics	Intervention Group (*N* = 487)		Comparison Group (*N* = 487)	
	Yes	50	10%	49	10%
Drug Use Before Pregnancy		*n*	%	*n*	%
	No	138	30%	131	29%
	Yes	325	70%	327	71%
Birth Status of Target Child		*n*	%	*n*	%
	Not First Birth	281	58%	270	56%
	First Birth	204	42%	214	44%
Intimate Partner Violence Status		*n*	%	*n*	%
	No Intimate Partner Violence	330	71%	334	71%
	Intimate Partner Violence	135	29%	137	29%
Edinburgh Postnatal Depression Scale		*n*	*M (SD)*	*n*	*M (SD)*
	Initial EPDS Scores	482	5.42 (4.93)	484	5.51 (4.90)
Baseline HFPI Score[1]		*n*	*M (SD)*	*n*	*M (SD)*
	3-Month Total HFPI Score	457	229.84 (91.44)	456	231.56 (90.49)
Baseline HOME Score[1]		*n*	*M (SD)*	*n*	*M (SD)*
	2-Month Total HOME Score	448	27.97 (13.91)	439	28.94 (14.17)

[1] Note: The baseline equivalence requirement for HomVEE's (2014) moderate study rating (i.e., non-significant differences between groups on race/ethnicity, socioeconomic status, and baseline measures of outcomes) was met.

C. Comparison Group

Research Question 3

Demographic characteristics for comparison groups were reported in the prior section. Because comparison families were identified from the population of non-MIECHV-funded families participating in the HFI program, no recruitment or retention issues specific to these families were experienced. Non-MIECHV-funded families were recruited and retained by HFI staff as part of normal service delivery. All historical data were collected as described in *VIII.A Data Collection* and were available to the evaluation team from the database vendor. To control selection bias, propensity score matching was utilized, and balanced samples were created through this process. Strategies were in place to address baseline equivalence and contamination; see *IX.C Design Specific Components*.

D. Discussion and Interpretation

The FY2018 evaluation allowed Indiana to build on findings from prior evaluations (particularly FY2016) to examine the role that fidelity plays in MHC from both an implementation evaluation (RQ1) and outcome evaluation (RQ2 & RQ3) perspective. Through this design, each evaluation question incorporated fidelity as a lens through which data were analyzed and results were interpreted. When examined collectively, the results reinforce the salience of implementation fidelity within the MHC model. Moreover, findings are consistent with the literature, which has shown that fidelity may moderate the effects of interventions on outcomes (Carroll et al., 2007). Specifically, the current study demonstrated that increased fidelity was associated with improved family and staff outcomes. It also identified the components of the model that, when implemented with fidelity, were the most predictive of outcomes for families and home visitors. These findings emphasize the importance of a focus on multiple dimensions of fidelity when implementing the enhancement. Furthermore, the results may isolate unique pathways through which home visitors and families benefit from different aspects of the model. However, it is important to note that given the small effect sizes observed for all family outcomes, the

magnitude of relationships was very limited. Specifically, increased fidelity to the treatment accounted for approximately 1% to 2% of the variability in the outcomes. Along with providing an increased understanding of the fidelity-outcomes link within MHC, the evaluation identified staff perceptions of current implementation gaps, existing supports and barriers, as well as additional resources that would support improved implementation. Results from each evaluation question are discussed in the following sections.

Research Question 1

RQ1 examined the supporting factors and barriers associated with implementing MHC with fidelity.

Supporting Factors. Most respondents agreed that as of summer 2019, the majority of fidelity criteria were being met; however, they noted that events closely preceding the interviews (e.g., May 2019 MHC Training, FY2016 evaluation reporting) had spurred improvements, which were still in the early stages of development. Sites' ability to implement MHC was supported by a variety of factors. First, respondents reported that because MHC was strongly aligned with home visiting work, sites intentionally leveraged the enhancement to improve their work with families. Rather than treat it as an add-on or a requirement, sites viewed MHC as an integral part of HFI services. There was a shared understanding within sites about its importance, which promoted buy-in from staff, including site leadership and home visitors. Secondly, as an offshoot of the aforementioned alignment and shared understanding, agency leaders bolstered MHC by integrating the enhancement into all agency activities, communicating the importance of MHC, ensuring that time and resources were appropriately allocated, and involving mental health consultants in meetings and other collaborative work. Buy-in from leadership provided reinforcement for other staff. Third, collaboration among mental health consultants promoted improved fidelity by sharing forms for documentation, resources, strategies for supporting home visitors, general information about MHC, and best practices for meeting model expectations. Fourth, mental health consultant's expertise/experience supported the implementation of MHC because mental health consultants' backgrounds were highly aligned with the duties and tasks necessary to implement MHC model expectations with fidelity (e.g., review families, provide consultation and reflective supervision, and design and deliver training efficiently and with high quality). Finally, respondents noted the value of MHC-specific training, and they specifically highlighted the MHC training provided in May 2019, which included reviews of the model expectations and associated documentation, question and answer opportunities, and resources.

Deviations from the Model. According to respondents, two main fidelity concerns existed for sites: documentation and monthly reviews of all MIECHV families. *Documentation.* Staff attributed some challenges with documentation to a lack of clarity related to expectations (e.g., what activities should be documented, how secondary activities should be documented) and limited experience with the data system. Moreover, respondents noted that longstanding documentation issues were compounded by the spring 2019 database transition, which created technical challenges and a learning curve for staff. Additional concerns included difficulty contacting technical support staff and long delays for support tickets. Documentation issues are a major concern because (as noted by mental health consultants) secondary activity data are the primary tool for reporting the extent to which model components have been completed; therefore, documentation inaccuracies affect the extent to which all components of the model are

accurately captured. Moreover, they are a primary means for assessing MHC fidelity. Another major concern is that documentation challenges are relatively widespread. Over half (57%) of respondents noted challenges completing documentation, which suggests that most site struggle with this model expectation. *Monthly Family Reviews.* Most (79%) mental health consultants and program managers reported that all MIECHV-funded families were being reviewed monthly at the time of the interview (August-September 2019); however, participants noted that in most sites, only high priority families had been reviewed and documented monthly prior to a training in May 2019 that clarified expectations. As with documentation issues, family review inconsistencies had been widespread; 86% (6/7) of mental health consultants said that they had incorrectly completed and/or documented family reviews before May 2019. Unfortunately, even after receiving clarification, some sites (particularly sites with larger caseloads) continue to struggle to review all families each month. Deviations from the model in this area are especially problematic in light of family outcome results from FY2018 that show increased fidelity of family reviews at both the site-level and family level significantly predicted improved outcomes for families. Other related challenges included the need to prioritize family reviews in light of other model expectations (e.g., clinical risk assignment, training) and requests from mental health consultants for additional clarity and guidance related to the expectation.

Barriers. As noted above, the 2019 data system migration and a lack of clarity related to model expectations hindered most sites from implementing MHC in adherence to all fidelity criteria. Respondents described a lack of access to the new data system (e.g., unavailable login credentials, user privileges), difficulty learning to use the new system and accessing technical support, limited features to support MHC (e.g., reports showing progress toward MHC model expectations), and poor functionality (e.g., difficult navigation, missing data). When discussing challenges related to meeting model expectations, interviewees sought additional clarity related to family reviews and documentation expectations.

Other barriers included caseloads, mental health consultants' time capacity, scheduling challenges, and home visitors' buy-in. First, according to participants, large caseloads made it difficult for sites to meet the monthly family review and clinical consultation expectations because these expectations require reviews/discussions of all MIECHV-funded families. Moreover, larger caseloads were perceived to negatively affect the quality of MHC delivery because less time could be devoted to each family. Secondly, mental health consultants reported that in some sites, an inadequate amount of time was dedicated to clinical consultation and reflective practice. Third, given the unpredictability of schedules, canceled clinical consultation and reflective practice sessions are often difficult to reschedule. Finally, program managers noted that some home visitors see MHC as an extra requirement and fail to see the value in participating in the extra support. As a result, some may be hesitant to participate fully in reflective practice and clinical consultation (i.e., lower levels of participant responsiveness). Interestingly, this perception was not shared by mental health consultations who generally reported buy-in from home visitors.

Additional Resources. Through the interviews, program managers and mental health consultants identified several resources that would support increased fidelity for the MHC enhancement. As noted above, collaboration among mental health consultants was identified as a key support, and respondents requested additional opportunities for collaboration including more

frequent in-person meetings and additional reflective practice for mental health consultants. Staff also saw value in training for mental health consultants that was specific to their duties through MHC, especially for reflective supervision. Respondents noted that general trainings (including existing reflective supervision trainings) were not well aligned to mental health consultants' roles within MHC and were not geared toward individuals with advanced degrees/licenses and/or clinical experience. Finally, respondents indicated that more detailed documentation guidelines (e.g., standardized forms, additional clarification of and rationale for model expectations, documentation reviews/audits, and detailed roles and responsibilities) would help to alleviate existing confusion and improve the quality of documentation.

Summary. In summary, staff reported that MIECHV sites have improved the extent to which MHC is implemented with fidelity across the majority of model expectations; however, deficiencies in the areas of documentation and monthly family reviews create notable challenges because of how these fidelity criteria influence overall reporting and family outcomes, respectively. While a variety of barriers and challenges exist, staff believe that with additional opportunities for professional development and collaboration, implementation fidelity can be improved. Because these responses were drawn from nearly all program managers and mental health consultants implementing MHC, the results were generalizable to Indiana's MIECHV sites.

Research Question 2

The results from research question 2 demonstrated the importance of fidelity within the context of MHC. Specifically, this research question demonstrated that specific dimensions of fidelity make at least partially unique contributions to outcomes for home visitors.

First, there was a large, significant relationship between delivery quality and perceived quality of the resources. Specifically, this relationship suggested that as 1) reflective processes and skills (e.g., integrating emotion and reason, talking about emotions), 2) mentoring (e.g., trusting relationships, engagement, time to come to own solutions), 3) supervision structure (e.g., consistent schedule), and 4) mentalization (e.g., explores cultural considerations, incorporates families' unique experiences) increased within MHC activities, the quality, relevance, and timeliness of MHC resources increased for home visitors.

Secondly, increased participant responsiveness was associated with greater self-efficacy and greater compassion satisfaction among the home visitors. Specifically, this relationship suggested that when home visitors were more responsive to the enhancement (e.g., increased confidence when discussing interactions and relationships between parents and young children or when examining thoughts, feelings, strengths, and growth areas), they had higher levels of self-efficacy (e.g., increased confidence addressing alcohol/drugs, mental health, partner violence, behavior management, and development with families) and compassion satisfaction (i.e., job satisfaction related to their role as a home visitor).

Finally, an increase in adherence to the structural aspects of fidelity (exposure and adherence) was associated with an increase in burnout (a small effect), which suggested that home visitors may find participating in MHC be demanding and burdensome when it is implemented with fidelity to the model.

Summary. As noted above, the results reinforce the integration of a multi-dimensional understanding of fidelity in the implementation of MHC. Specifically, delivery quality, participant responsiveness, and structural fidelity appear to make unique contribution to staff outcomes. In practice, attention to and quality in each of these areas of fidelity should be associated with better outcomes for staff. Additionally, these findings may provide insight into the unique pathways through which home visitors benefit from participation in MHC. Specifically, the results may suggest that home visitors benefit most from the reflective/relational components of MHC. This interpretation is consistent with the MHC theory of change, which states that improved outcomes occur within the context of the relationships that consultants build with home visitors (Hunter et al., 2016; Johnston & Brinamen, 2012; Watson, et al., 2016). Given the sample size ($N = 66$) and response rate (80.4%), these results are likely generalizable to the population of home visitors participating in MHC.

Research Question 3

In research question 3, the study examined the relationship between participation in the MHC model and family outcomes using a dose-response framework that considered increased fidelity to the treatment model as an increased dose. By accounting for fidelity in the study, the evaluation contributed to a more robust understanding of MHC and how it can benefit families.

Preliminary steps taken to prepare the fidelity score provided valuable insight into the individual MHC model components that are positively related to improved outcomes for families. Specifically, preliminary evidence points to possible benefits specifically associated with family reviews, reflective practice, clinical consultation, clinical risk review, and MHC training. Moreover, there appears to be a link between longer family participation in MHC and improved outcomes. In practice, an increased focus on fidelity in these areas may support improved outcomes for families; however, it is important to note that the effect sizes were very small, which suggest that the magnitude of any improvements in family outcomes could be limited.

RQ3a. The results demonstrated that fidelity to the model predicted family outcomes within the MHC treatment group for the HOME outcomes, but not for the HFPI outcomes. For the HOME subscales, there were effects of total fidelity score on the slopes of outcome measure, which indicated greater improvement in the outcome as fidelity increased. For the HFPI subscales, no effects of fidelity on change were observed. For the HOME, these results suggested that within the dose-response framework utilized for the evaluation, increased fidelity was associated with improved outcomes; however, the small effect sizes indicated that the magnitude of the relationships were limited.

RQ3b. On the HOME subscales, all groups tended to improve over time except the Low Fidelity MHC treatment group, which either did not improve over time or demonstrated less improvement than other groups. On the HFPI subscales, the data suggested no changes over time for all groups except the Low Fidelity treatment group, which actually worsened over time on these subscales. Two possible interpretations have emerged from these results. 1) Because the majority of the total fidelity score was determined by factors occurring at the site level rather than at the family level, it is possible that poor performing sites both have difficulty adhering to the MHC model with fidelity and execute their invention more poorly. 2) Within the dose-

response framework used for the evaluation, increased fidelity to the treatment model is interpreted as an increased treatment dose. Therefore, an alternative possibility is that because MIECHV sites are specifically selected due to their location in high-risk counties, low fidelity MHC families may represent higher-risk families that essentially did not receive a treatment (based on the theorized dose-response relationship that defines lower fidelity as a lower treatment dose). This may suggest that families from high risk communities not receiving a treatment, or receiving a low-quality treatment, may perform more poorly in general than those from lower risk communities. In this case, the fact that the high fidelity MHC treatment group performed the same as its matched-comparison group may indicate that it is actually performing much better than it otherwise would have without the treatment because it is composed of families from high-risk counties.

Summary. The results of RQ3 reinforce the importance of fidelity when implementing MHC. These findings are consistent with the literature, which has shown that fidelity moderates the effects of interventions on outcomes (Carroll et al., 2007). As such, improving fidelity of implementation should contribute to improved outcomes for families enrolled in the program. It is important to note that there were small effect sizes observed for all family outcomes, and increased fidelity accounted for approximately 1% to 2% of the variability in the outcomes. While this does not diminish the importance of fidelity, it does suggest that the magnitude of improvements may be very limited as MHC fidelity is increased. Additionally, the results suggest that home visitors and families benefit from different fidelity criteria, which may elucidate the pathways through which MHC supports improved outcomes for key stakeholders. As shown in RQ2, home visitors appear to benefit most from increased fidelity to reflective/relational components (e.g., MHC delivery quality and participant responsiveness). Families, on the other hand, appear to benefit most from increased fidelity to model adherence (e.g., completion of clinical risk assignment, monthly family review, reflective practice, clinical consultation, and training requirements) and exposure components (e.g., amount of time that families spend in the program). However, it is important to note that the theory of change suggests that families benefit from the relational components of MHC when home visitors replicate the supervisor-supervisee relationship in their work with families (i.e., through parallel process) (see *VII.E Theory of Change*). Because these analyses were exploratory, it is possible that the design was not adequate to detect family outcomes driven by parallel process. Additional research and evaluation are needed to understand fully the relationships among fidelity criteria and outcomes, as well as how these relationships may differ for stakeholder groups.

Unintended Findings

RQ1 – Home Visitor Buy-in. While the majority of program mangers identified home visitor buy-in as a barrier for MHC implementation, this concern was not shared by mental health consultants (Mental Health Consultants: 14%; Program Mangers: 71%). Based on the interview responses, it was unclear why the two groups perceive home visitors' buy-in in MHC differently.

RQ1 – The Presence of Supervisors during Home Visitors' Reflective Practice. While it did not reach the 40% threshold for inclusion in the main results section, the majority of mental health consultants (57%) expressed concerns about including supervisors in reflective practice provided to home visitors. Specifically, they noted that with the supervisor present, some home visitors were less willing to discuss challenges, which limited the efficacy of reflective practice.

Overall, program managers did not share this concern (or were not aware of this issue); however, one program manager noted that his or her site offered reflective practice without the supervisor present and indicated that MHC had benefited from separating reflective practice from supervision. Given its importance for mental health consultants, the supervisor's role in reflective practice should be examined further.

RQ3 – HFPI – Depression. Throughout the analyses for research question 3, the *Depression* subscale from the HFPI behaved inconsistently compared to other scales, the theory of change, and the expectations of the evaluation team. In Step 1 of the analysis, this scale was shown to be negatively correlated with home visitor ratings of structural fidelity. Furthermore, in the matched groups comparisons, the slopes of the High Fidelity MHC group (b = -0.002, se = .01, p = .982) and High Fidelity Comparison group (b = -0.05, se = .01, p < .001) were both negative. This negative slope suggested that depression for the High Fidelity Comparison group worsened over time. It is unclear why there was a paradoxical change in depression scores over time for these groups. Before conclusions may be drawn about this relationship, additional evaluation is recommended (see *XII.G Recommendations*).

E. Limitations

RQ1. Evaluation data were drawn from self-report interviews, which can create a number of limitations including, but not limited to, social desirability bias, attribution issues, and memory errors. Moreover, responses reflected the perceptions of the interviewee and may not be independently verifiable. Finally, responses reflect each participant's understanding of specific MHC model expectations and may have varied across respondents and/or have been inconsistent with the MHC contractual expectations. As applicable, efforts have been made throughout this report to triangulate interview responses with other data sources for the most accurate interpretations. **RQ2.** Due to the small sample size, the analysis was "rank deficient" (i.e., there were more measurement items (88) than observations (N = 66)). As a result, the overall analysis is not as reliable as it would be were it not rank deficient; therefore, some caution in interpretation is necessary. In particular, one should use caution when interpreting any small effects (both their size and direction) because small effects in a small sample may not be present or even reverse direction in a larger sample. Large effects may be viewed as preliminary evidence of relationships among variables; however, these findings would benefit from further investigation either via replication or qualitative methods. **RQ3.** *Propensity Score Matching/Quasi-Experimental Research.* It should be noted that while propensity score matching was used to create comparison groups that were similar to the families participating in MHC, the process cannot control all bias and should not be considered equivalent to a true experimental study. The analyses may be limited by the existence of variables that predict family participation and/or outcomes but were not available to the evaluation team. These analyses should be interpreted as only preliminary evidence of program impacts (Sommers et al., 2013). *HFPI.* Overall, the HFPI appeared less sensitive to detecting changes in family outcomes over time. Currently, it is unclear if these issues are the result of the outcomes themselves (i.e., the types of outcomes the HFPI is designed to detect are particularly resistant to change) or measurement problems (i.e., issues with the scale itself or how it is administered or scored). *Fidelity.* As noted in RQ1, while fidelity has improved since the Localized MHC model roll out in 2016, opportunities for sites to improve the implementation of MHC remain, especially in the areas of documentation and family review. It is important to interpret the findings within the context of

Indiana's ongoing efforts to improve fidelity. Specifically, while families within the High Fidelity MHC treatment group had higher fidelity scores than families in the Low Fidelity MHC group, it is likely that many of the High Fidelity MHC families received services that did not satisfy model requirements. Moreover, because *a priori* knowledge was limited, a median split was used to determine placement into the high and low fidelity treatment groups. Future research should identify high and low fidelity families based on the clusters of fidelity scores (those below -1, those clustering around 0, and those above 1) identified in the current study (see RQ3a). *Home Visitor Fidelity Measures.* Structural fidelity as measured by the home visitor survey (IN MHC Fidelity Scale) was not associated with structural fidelity as measured by the secondary activity data. Additional review of these measures suggested that home visitors may not the best source of information for some aspects of structural fidelity. While they should have strong insights related to the model expectations in which they are directly involved (e.g., clinical consultation, reflective practice, and training), accurately recalling the extent to which all of their families were discussed in clinical consultation or how often training was offered over the course of a year might be difficult. Moreover, home visitors may be less attuned to clinical risk assignment and monthly family reviews because they are not directly involved in these processes. Finally, psychometric concerns emerged related to the IN MHC Fidelity Scale. The development of new tools to assess fidelity should be considered as part of ongoing fidelity improvement.

F. Design Specific Components

Controlling Confounding Factors. Findings from home visiting (and related) evaluation/research have identified a variety of factors that are associated with family outcomes and participation in home visiting interventions (e.g., Boller et al., 2014; CDC, 2020; Daro et al., 2003; MIECHV, 2016). This research was consulted during the development of the study to identify and plan for confounding factors. Resulting variables included caregiver demographics (education, race, age, socioeconomic status, language, urbanicity, substance use, criminality history), family risk (intimate partner violence, family stress, parenting stress, childhood history of care, emotional functioning, attitudes towards and perception of children, discipline of children), mental health, and baseline parenting efficacy. Propensity score matching was employed to control confounding factors. Propensity score matching has been identified as an effective method for addressing confounding factors and reducing bias (Brookhart et al., 2013; Kahlert et al., 2016; Rosenbaum & Rubin, 1983; Schroeder et al., 2016). The literature suggests that when the process has been implemented successfully, "treated and untreated [individuals] tend to have the same distribution of measured confounders, something that we would also achieve using randomization" (Brookhart et al., 2013, p. 2). To control bias effectively, the evaluation team sought to identify comparison groups that contained 1) prescreened individuals with motivation and incentives (or deterrents) to participate that are similar to those of the treatment group, 2) individuals from close geographical proximity to the treatment group (e.g., regional), and 3) those who have similar pretest scores on the outcome of interest compared to the treatment group (Cook et al., 2008; Glazerman, et al., 2003; Sommers et al., 2013). The FY2018 study was designed to adhere to HomVEE's standards for a moderate study rating (the highest possible for a matched comparison group design). Specifically, 1) potential confounding factors have been minimized as recommended by HomVEE (i.e., more than two subjects/participants were represented in the treatment and comparison groups, and data were uniformly collected across all sites) and 2) both baseline demographic characteristics (e.g.,

race/ethnicity, socioeconomic status) and baseline outcome measures were incorporated into the propensity score (HomVEE, 2014, 2017). Propensity score matching yielded balanced samples based on the covariates of interest, and multivariate and univariate tests revealed no evidence of imbalance, including on the covariates identified by HomVEE (2014) for a moderate study rating: race/ethnicity, socioeconomic status, and baseline outcomes. See *XI.B Sample and Comparison Group* for a detailed discussion of baseline equivalence.

XII. Evaluation Successes and Challenges, Conclusions, and Dissemination of Evaluation Findings

A. Strategies That Facilitated Evaluation Implementation

The evaluation was implemented as a partnership between Diehl Consulting Group and state-level DCS Prevention Team, with support from site managers, other program staff, and the database provider. The FY2018 evaluation utilized the DCS Prevention Team as the conduit through which evaluation instruments were administered to home visitors. DCS Prevention Team staff sent an advance email and recruitment email (with links) to site staff for each evaluation activity that provided an overview of the activity/timing and its importance for HFI. Additionally, response rates were reviewed with program managers by DCS Prevention Team staff during data collection windows to identify participation barriers and other issues. These procedures increased buy-in and participation from staff. During February 2019, members of the evaluation team reviewed FY2016 findings with site staff (including program managers and mental health consultants) and provided a preview of the preliminary FY2018 data collection timeline. Opportunities were provided for questions and feedback. The evaluation utilized secondary activity and family data housed in the Enlite system. Exports of required data were readily available from the database provider; state HFI leadership facilitated interactions between the evaluation team and the database vendor and provided contextual support for data requests. The evaluation team provides regular updates to the Indiana Home Visiting Advisory Board (INHVAB). Finally, state-level leadership are heavily involved in the evaluation process and meet with the evaluation team at least monthly.

B. Evaluation Successes

The FY2018 evaluation provided Indiana with the opportunity to explore MHC implementation and outcomes through the lens of fidelity. By accounting for fidelity in the study, nuanced evaluation questions were proposed that contributed to a more robust understanding of MHC and its benefits for families and staff. The FY2018 evaluation incorporated more rigorous statistical methods than prior evaluations, which increased the validity of the findings and the resulting conclusions. Using more rigorous methods identified relationships that were overlooked in prior evaluations, which contributed to an increased understanding of how MHC benefits stakeholders. While findings generally confirmed the MHC theory of change, they also challenged existing assumptions (e.g., the relationships between individual fidelity dimensions and outcomes). In practice, these findings will help Indiana to understand more specifically how MHC can benefit stakeholders, supports and barriers influence the implementation, and resources could support model improvements.

A second success was that strategies to increase survey responses rates for home visitors were fruitful during the FY2018 evaluation. Specifically, the response rate for the home visitor survey in FY2018 was 80.4%, compared to 63.6% for a similar home visitor survey in FY2016.

C. Challenges
During spring 2019, Indiana transitioned to a new data system for HFI. While the evaluation team was not directly affected by this change, frontline staff noted a variety of challenges collecting and entering data during the transition. Some data integrity issues were identified in the weeks and months immediately following the transition. Moreover, additional time devoted to data collection affected staff capacity in other areas. To mitigate these challenges, the evaluation timeline was designed to minimize data collection activities during the transition window. Specifically, 1) staff surveys and interviews were completed in late summer 2019 to accommodate staff capacity concerns, and 2) data requests were completed either a) prior to the transition (in collaboration with the previous vendor) or b) during winter 2020 (in collaboration with the current vendor). These strategies lessened burdens on staff and reduced the risk of inaccurate or incomplete data. Because all data collection had occurred by December 31, 2019, the COVID-19 pandemic created no challenges for the FY2018 evaluation.

D. Adherence to Proposed Plan
The evaluation was completed in adherence to the approved evaluation plan. No deviations were made to the design, evaluation questions, data collection instruments or procedures, or timeline without approval from HRSA.

E. Key Findings
RQ1. Overall, mental health consultants and program managers reported that the majority of model expectations were met; however, responses suggested that most sites experienced challenges meeting documentation expectations and that some sites struggled to complete monthly reviews for all MIECHV-funded families. Furthermore, the majority of sites noted that monthly reviews were not completed and/or documented in adherence to model expectations prior to May 2019. When asked to describe implementation supports currently available, participants identified MHC alignment with existing work (93%), agency/management support for MHC (86%), collaboration among mental health consultants (57%), mental health consultants' expertise/experience (50%), and MHC training provided in May 2019 (43%). Analysis of interview responses revealed three additional supports that would benefit implementation: increased collaboration among mental health consultants (50%), training for mental health consultants that focuses specifically on their MHC responsibilities (50%), and additional documentation guidelines (43%). Key barriers impeding implementation fidelity included spring 2019 data system migration (86%), model expectation clarity (79%), caseloads (57%), mental health consultants' time capacity (57%), scheduling challenges (43%), and home visitors' buy-in (43%).
RQ2. A fully linked model was created that explored the unique relationships between each fidelity measure and each staff outcome. The model suggested that the individual aspects of fidelity make different, at least partially unique contributions to the home visitor outcomes. There was a large, significant relationship between delivery quality and perceived quality of the resources; as ratings of the quality of delivery increased, increases were observed in perceived quality of resources. Increased participant responsiveness (i.e., confidence completing key aspects of MHC) was associated with greater self-efficacy (a medium-sized effect) and greater compassion satisfaction (a small effect) among the home visitors. An increase in adherence to the structural aspects of the MHC model was associated with an increase in burnout (a small

effect). This latter relationship suggests that home visitors may find participating in MHC be demanding and burdensome when it is implemented with fidelity; however, because this was a small effect, these results suggest a need to further explore the obstacles home visitors may encounter.

RQ3a. The results suggested that fidelity to the MHC model predicted family outcomes within the MHC treatment group for the HOME outcomes, but not for the HFPI outcomes. For the HOME subscales, there were effects of total fidelity score on the slopes of outcome measure, which indicated greater improvement in the outcome as fidelity increased. These effects were statistically significant in the ordinary least squares (OLS) regression for all of the HOME measures, but not always so when the random effect of site was also included in the model. In both cases, effects were very small, with fidelity accounting for 1% of the variance in the change in the HOME outcomes over time. For the HFPI subscales, the results of the mixed linear models (MLM) and OLS regressions were equivalent, and no effects of fidelity on the change in these subscales were observed. The main MLM and OLS analyses were supplemented by exploratory analyses that compared the mean change over time in each of six subscales for two groups composed of MHC families with the highest and lowest fidelity scores. These analyses confirmed the findings from the primary analyses. In summary, RQ3a results suggest an effect of fidelity to the MIECHV treatment model on HOME outcomes, albeit a small one.

RQ3b. On the HOME subscales, all groups tended to improve over time except the Low Fidelity MHC treatment group, which either did not improve over time or demonstrated less improvement than other groups. On the HFPI subscales, the data suggested no changes over time for all groups except the Low Fidelity treatment group, which actually worsened over time on these subscales. Two possible interpretations have emerged from these results. 1) Because the majority of the total fidelity score was determined by factors occurring at the site level rather than at the family level, it is possible that poor performing sites both have difficulty adhering to the MHC model with fidelity and execute their invention more poorly. 2) An alternative possibility is that because MIECHV sites are specifically selected due to their location in high-risk counties, low fidelity MHC families may represent higher-risk families that essentially did not receive a treatment (based on the theorized dose-response relationship that defines lower fidelity as a lower treatment dose). This may suggest that families from high risk communities not receiving a treatment, or receiving a low-quality treatment, may perform more poorly in general than those from lower risk communities. In this case, the fact that the high fidelity MHC treatment group performed the same as its matched-comparison group may indicate that it is actually performing much better than it otherwise would have without the treatment because it is composed of families from high-risk counties.

F. Implications of Findings

The results of this evaluation suggest that participation in MHC may provide benefits for families, home visitors, and agencies; however, implementing the enhancement with fidelity to the model is essential for improving outcomes for stakeholders. For families in particular (RQ3), the evaluation found that increased fidelity to the treatment model was associated with improved outcomes in the areas of *emotional/verbal responsivity* ("the communicative and affective interactions between the caregiver and the child"), *organization of physical and temporal environment* ("how the child's time is organized outside the family house, [and] what the child's personal space looks like), and *parent involvement* ("how the adult interacts physically with the child") (Totsika & Sylva, 2004, p. 26). Moreover, matched-comparison analyses may suggest

that when implemented with higher levels of fidelity, MHC could provide some mitigation for negative parenting outcomes experienced by families in high-risk communities. The evaluation also provided preliminary evidence that identified the model components that appear to have the strongest relationships with family outcomes when implemented with fidelity. Specifically, preliminary evidence points to possible benefits specifically associated with family reviews, reflective practice, clinical consultation, clinical risk assignment, and MHC training. Moreover, there appears to be a link between longer family participation in MHC and improved outcomes. However, it is important to note that given the small effect sizes observed for all family outcomes, the magnitude of family improvements may be very limited as fidelity is increased. For home visitors (RQ2), the study identified how individual dimensions of MHC fidelity contribute to better outcomes. Specifically, the quality of MHC delivery was positively associated with perceived *quality of resources*. Increased home visitor responsiveness (i.e., confidence completing key aspects of MHC) was associated with greater *self-efficacy* (i.e., confidence providing support in the areas of drugs/alcohol, mental health, partner violence, behavior management, and child development) and greater *compassion satisfaction* (i.e., job satisfaction related to helping others) among the home visitors. While the effect size was small, there was some evidence to suggest that when MHC is implemented with higher levels of structural fidelity (i.e., model adherence and exposure), there was greater *burnout* among home visitors, and this finding is important to consider as adaptations to the model are considered. At the program level (RQ1), qualitative interview responses indicated that while improvements to implementation fidelity have occurred, some sites continue to struggle with documentation and monthly reviews for all MIECHV-funded families. Moreover, at the time of the interviews, strategies to maximize fidelity were still in the early stages of implementation. Mental health consultants and HFI program managers identified a variety of factors that are currently in place that support the implementation of MHC, as well as barriers that hinder implementation with fidelity. Finally, respondents provided recommendations for improving the quality of implementation. In summary, findings examining family and home visitor outcomes (RQ3 and RQ2, respectively) provide preliminary support for continuation of the model, especially with strategies in place to improve the fidelity of implementation. Interview responses (RQ1) provide valuable insight to understanding deficiencies in current implementation, identifying strengths and gaps, and developing strategies to strength the model and its implementation. These data have and will continue to inform program-level decision-making in Indiana. Preliminary findings from the FY2018 evaluation were presented to DCS Prevention Staff during spring 2020, and these data were used to guide the development of the Indiana's FY2020 MIECHV application, particularly through strategies to improve MHC and increase fidelity across sites through a variety of new supports and resources. Moving forward, the data suggest that by increasing MHC fidelity across participating sites, Indiana may experience improved outcomes for families participating in the program; however, given the small effect sizes observed in the FY2018 evaluation, the magnitude of family improvements may be very small as fidelity is increased.

G. Recommendations

Across all research questions, the findings emphasize the importance of fidelity for improving outcomes for families and home visitors. Greater fidelity to the model is associated with better outcomes, which indicates that continuing to emphasize improved fidelity is important for the implementation teams at the state level. Grounded in the FY2018 evaluation findings, the

following recommendations provide guidance for improving fidelity, measurement, and evaluation.

1. Ongoing Training and Support for MHC Implementation: The findings suggest that sites would benefit from additional training and support that are directly related to completing and appropriately documenting the MHC model expectations. In particular, the data suggest that Indiana should consider training that covers multiple aspects of implementation fidelity, including model adherence, exposure, participant responsiveness, and delivery quality. As applicable, recommendations from mental health consultants and program managers should be considered. These included increased opportunities for mental health consultants to collaborate, training for mental health consultants that is specific to their MHC role (including reflective practice), and detailed guidelines for documentation.

2. Ongoing Fidelity Monitoring: Consideration may be given to developing strategies to monitor implementation fidelity on an ongoing basis. Doing so would allow leadership to identify implementation issues and make real-time course corrections. Implementation fidelity is improved when program components are defined *a priori* and monitored for compliance (Mihalic, 2004).

3. Review and Revision of Model Expectations: The results have provided initial evidence of the extent to which individual model expectations are related to staff and family outcomes. Where applicable, existing fidelity criteria should be reviewed and revised to increase the focus on the model components that have the strongest relationships with program outcomes.

4. Develop and/or Refine Fidelity Instruments. As noted in *XI. E Limitations*, concerns were noted related to the *IN MHC Fidelity Scale*, which was developed as part of the FY2018 evaluation. Valid and reliable measures for assessing MHC fidelity are essential for monitoring compliance and ultimately, improving fidelity (Prinz & Moncher, 1991). Therefore, the development of new tools to assess fidelity should be considered as part of ongoing fidelity improvement.

5. Further Evaluation and Research. The results of this evaluation provided new insight into the role of fidelity in moderating family outcomes and how different stakeholders could benefit from participating in MHC. As such, further evaluation and research should utilize designs that allow outcomes to be examined within the context of fidelity. Given the implementation improvements that mental health consultants and program mangers described, future evaluation should include a focus on MHC implementation after May 2019. Finally, future evaluation should incorporate examination of the Depression subscale from the HFPI. As noted above (*XI.D. Discussion and Interpretation: Unintended Findings*), results were inconsistent with the theory of change, and additional research is recommended before conclusions can be drawn about the relationship between MHC and depression.

H. Detailed Plan for Dissemination

The dissemination plan will include opportunities for sharing lessons learned to all MIECHV recipients and to the home visiting field broadly. Locally, interim reports will be shared with the early childhood system stakeholders as part of the dissemination process within the state to help identify gaps in services, particularly for these primary needs. Local stakeholders include implementing agencies and the Indiana Home Visiting Advisory Board (INHVAB). Sharing opportunities include MIECHV and home visiting related meetings, conferences, and the MIECHV website. Additionally, these reports will be shared with the Federal program staff and

technical assistance staff, as the State MIECHV team provides updates on the progress of the grant. Finally, conference presentations, including the Institute for Strengthening Families, will be considered during the project.

References

Anseel, F., Lievens, F., Schollaert, E., & Choragwicka, B. (2010). Response rates in organizational science, 1995-2008: A meta-analytic review and guidelines for survey researchers, *Journal of Business and Psychology, 25*(3), 335-349.

American Educational Research Association (AERA), American Psychological Association (APA), & National Council on Measurement in Education (NCME) (2014). *The standards for educational and psychological testing.* Washington DC: American Psychological Association.

Avellar, S., & Paulsell, D. (2011). *Lessons learned from the home visiting evidence of effectiveness review.* Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Austin, P.C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research, 46*(3), 399-424.

Barnard, K.E., Booth, C.L., Mitchell, S., & Telzrow, R.W. (1988). Newborn nursing models: A test of early intervention to high-risk infants and families. In E.G. Hibbs (Ed.), *Children and families: Studies in prevention and intervention* (pp. 63–81). Madison, CT: International Universities Press.

Blundell, R., Dearden, L., & Siansi, B. (2005). Evaluating the effect of education on earnings: models, methods and results from the National Child Development Survey. *Journal of the Royal Statistical Society: Series A, 168*(3), 473-512.

Boller, K., Daro, D., Del Grosso, P. (2017) *Supporting Evidence-Based Home Visiting to Prevent Child Maltreatment.* Washington, DC: Children's Bureau, Administration for Children and Families, U.S. Department of Health and Human Services.

Bradley, R. H. (1993). Children's home environments, health, behavior, and intervention efforts: A review using the HOME inventory as a marker measure. *Genetic, Social and General Psychology Monographs, 119*, 437–490.

Breitenstein, S.M., Gross, D., Garvey, C., Hill, C., Fogg, L., & Resnick, B. (2010). Implementation fidelity in community-based interventions. *Research in Nursing & Health, 33*(2), 164-173.

Brookhart, M. A., Wyss, R., Layton, J. B., & Stürmer, T. (2013). Propensity score methods for confounding control in nonexperimental research. *Circulation, Cardiovascular Quality and Outcomes*, *6*(5), 604–611.

Burwick, A., Zaveri, H., Shang, L., Boller, K., Daro, D., & Strong, D. (2014). Costs of early childhood home visiting: An analysis of programs implemented in the supporting evidence-based home visiting to prevent child maltreatment initiative. Princeton, NJ: Mathematica Policy Research.

Caldwell, B. & Bradley, R. (1984). *Home Observation for Measurement of the Environment (HOME) - Revised Edition.* University of Arkansas, Little Rock.

Caldwell, B. & Bradley, R. (1984). The home inventory and family demographics. *Developmental Psychology, 20*(2), 315-320.

Caliendo, M. & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys, 22*(1), 31-72.

Carroll, C., Patterson, M., Wood, S., Booth, A., Rick, Jo, & Balain S. (2007). A conceptual framework for implementation fidelity. Implementation Science, *2*(40), 1-9.

Center for Disease Control (2020). Risk and Protective Factors. Atlanta: National Center for Injury Prevention and Control, Division of Violence Prevention Retrieved from https://www.cdc.gov/violenceprevention/childabuseandneglect/riskprotectivefactors.html

Center for Prevention Research and Development (2018). MIECHV home visiting and continuous quality improvement survey. Champaign, IL: Center for Prevention Research and Development at the University of Illinois.

Cherry, A. J. (2000). *A research primer for the helping professions: Methods, statistics, and writing*. Australia: Brooks/Cole.

Chin, W. W., Marcolin, B. L., & Newsted, P. R. (2003). A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study. *Information Systems Research, 14*(2), 189–217.

Cochran W.G. & Rubin D.B. (1973). Controlling bias in observational studies: A review. *The Indian Journal of Statistics, Series A., 35,* 417-466.

Coe, R. (2002). It's the effect size, stupid: What effect size is and why it is important. Annual Conference of the British Educational Research Association. University of Exeter, England. September 12, 2002.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New Jersey: Lawrence Erlbaum.

Cook, T.D., Shadish, W.R., & Wong, V.C. (2008). Three conditions under which experiments and observational studies produce comparable causal estimates: New findings from within-study comparisons. *Journal of Policy Analysis and Management,* 27(4), 724-750.

Cox, J.L. (1996). Perinatal mental disorders – a cultural approach. *International Review of Psychiatry, 8,* 9-16.

Cox, J.L., Holden, J.M., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. *British Journal of Psychiatry, 150*, 782-786.

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*, 16-29.

Cycyota, C.S. & Harrison, D.A. (2002) Enhancing survey response rates at the executive level: Are employee- or consumer-level techniques effective? *Journal of Management, 28*(2), 151-176.

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.

Datatude, Inc. (2016). Home Observation Measurement of Environment Inventory for Healthy Families Indiana FY 2015. Indianapolis, IN: Indiana Department of Child Services.

Department of Child Services, Healthy Families Indiana, & Datatude, Inc. (2016a). Evaluation Paper #2016-5 HOME Scale Part 3 of 3. Indianapolis, IN: Indiana Department of Child Services.

Department of Child Services, Healthy Families Indiana, & Datatude, Inc. (2016b). Evaluation Paper #2016-6 HOME Scale Part 3 of 3. Indianapolis, IN: Indiana Department of Child Services.

Department of Child Services, Healthy Families Indiana, & Datatude, Inc. (2016c). Evaluation Paper #2016-7 HOME Scale Part 3 of 3. Indianapolis, IN: Indiana Department of Child Services.

D'Agostino, R.B. (1998). Tutorial in biostatistics: Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine, 17*, 2265-2281.

Duran, F., Hepburn, K., Irvine, M., Kaufmann, R., Anthony, B., Horen, N., & Perry, D. (2009). *What works? A study of effective early childhood mental health consultation programs.* Washington DC: Georgetown University, Center for Child and Human Development.

Easton, K.L., Fry McComish, J., & Greenberg, R. (2000). Avoiding common pitfalls in qualitative data collection and transcription. *Qualitative Health Research, 10*(5), 703-707.

Eberhard-Gran, M., Eskild, A., Tambs, K., Schei, B., & Opjordsmoen, S. (2001). The Edinburgh Postnatal Depression Scale: Validation in a Norwegian community sample. *Nordic Journal of Psychiatry, 55,* 113-117.

Fan, W. & Yan, Z. (2010). Factors affecting response rates of the web survey: A systematic review, *Computers in Human Behavior*, *26*, 132-139.

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*, 175-191.

Fixsen, D.L., Naoom, S.F., Blasé, K.A., Friedman, R.M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature.* Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).

Fowler, F. J. (2004). Interview schedule. In M. S. Lewis-Beck, A. Bryman & T. Futing Liao (Eds.), *The SAGE encyclopedia of social science research methods*: Sage.

Gallen, Ash, Smith, Franco, & Willford (2016) How do I know that my supervision is reflective? Identifying factors and validity of Reflective Supervision Rating Scale, *Zero to Three, 37*(2), 30-37.

Galiana, L., Arena, F., Oliver, A., Sansó, N., & Benito, E. (2017) Compassion satisfaction, compassion fatigue, and burnout in Spain and Brazil: PRoQOL validation and cross-cultural diagnosis. *Journal of Pain and Symptom Management, 53*(3), 598-604.

Garvin, E. C., Cannuscio, C. C., & Branas, C. C. (2013). Greening vacant lots to reduce violent crime: A randomised controlled trial. *Injury Prevention, 19*(3), 198-203.

Glazebrook, C., Marlow, N., Israel, C., Croudace, T., Johnson, S., White, I.R., & Whitelaw, A. (2007). Randomized trial of a parenting intervention during neonatal intensive care. *Archives of Disease in Childhood. Fetal and Neonatal Edition, 92*(6), 438-443.

Glazerman, S., Levy, D.M., & Meyers, D. (2003). Nonexperimental versus experimental estimates of earnings impacts. *The Annals of the American Academy of Political and Social Science*, 589, 63-93.

Goedert, K.M., Boston, R.C., & Barrett, A.M. (2013). Advancing the science of neglect rehabilitation via improved statistical methods: Mixed linear modeling versus repeated-measures ANOVA. *Frontiers in Human Neuroscience, 7*, 211.

Green, S.B. & Salkind, N.J. (2011). *Using SPSS for Windows and Macintosh: Analyzing and understanding data* (6th edition). Boston: Prentice Hall.

Green, B.L., Everheart, M., Gordon, L., & Gettman, M.G. (2006). Characteristics of effective mental health consultation in early childhood settings. *Topics in Early Childhood Special Education, 26*(3), 142-152.

Green, B. L., Simpson, J., Everhart, M. C., Vale, E., & Gettman, M. G. (2004). Understanding integrated mental health services in Head Start: Staff perspectives on mental health consultation. *NHSA Dialog, 7*(1), 35-60.

Hansen, B. B. (2004), Full matching in an observational study of coaching for the SAT. *Journal of the American Statistical Association, 99*, 609-618.

Hansen, B. B., & Bowers, J. (2008). Covariate balance in simple, stratified and clustered comparative studies. *Statistical Science, 2*, 219-236.

Healthy Families Indiana. (2016) *Evaluation of the Indiana Maternal, Infant, Early Childhood Home Visiting (MIECHV) program: Healthy Families Indiana mental health consultation Program model enhancement: FY11 Evaluation – Study #2.* Indianapolis, IN: Indiana Department of Child Services.

Hsieh, Hsiu-Fang & Shannon, Sarah. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research, 15*(9), 1277-88. 10.1177/1049732305276687.

Ho D., Imai K., King G., Stuart E. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15*(3), 199–236.

Ho, D. E., Imai, K., King, G., & Stuart, E. A. (2011). MatchIt: Nonparametric preprocessing for parametric causal inference. *Journal of Statistical Software, 42*(8).

Iacus, S. M., King, G., & Porro, G. (2009). CEM: Software for coarsened exact matching. *Journal of Statistical Software, 30*, 1-27.

Iacus, S.M., King, G., & Porro, G. (2011). Multivariate matching methods that are monotonic imbalance bounding, *Journal of the American Statistical Association, 106*(493), 345-361.

Home Visiting Evidence of Effectiveness. (2017). Review process: Producing study ratings. Retrieved from http://homvee.acf.hhs.gov/Review-Process/4/Producing-Study-Ratings/19/5.

Home Visiting Evidence of Effectiveness. (2014). On equal footing: the importance of baseline equivalence in measuring program effectiveness. Washington, DC: Office of Planning, Research, and Evaluation. Retrieved from http://homvee.acf.hhs.gov/HomVEE_brief_2014-50.pdf.

Hunter, A., Davis, A., Perry, D.F., & Jones, W. (2016). *The Georgetown model of early childhood mental health consultation: For school-based settings.* Washington, DC: Center for Child and Human Development.

Johnson, K. & Brinamen, C.F. (2012). The consultation relationship – from transactional to transformative: Hypothesizing about the nature of change. *Infant Mental Health, 33*(3), 226-233.

Kahlert, J., Gribsholt, S. B., Gammelager, H., Dekkers, O. M., & Luta, G. (2017). Control of confounding in the analysis phase - an overview for clinicians. *Clinical epidemiology*, *9*, 195–204.

Kock, N., & Hadaya, P. (2018). Minimum sample size estimation in PLS-SEM: The inverse square root and gamma-exponential methods. *Information Systems Journal, 28*(1), 227–261.

Kormacher, J. (2000). The Kempe Family Stress Inventory: A review. *Child Abuse & Neglect, 24*(1), 129-140.

Koziol-McLain, J., Coates, C.J., Lowenstein, S.R. (2001). Predictive validity of a screen for partner violence against women. *American Journal of Preventive Medicine, 21*(2), 93-100.

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), 496-50.

LeCroy & Milligan Associates, Inc., (2004). *Healthy Families Parenting Inventory.*

LeCroy, C.W., Krysik, J., & Milligan, K. (2007). The Healthy Families Parenting Inventory. In J. Fisher & K. Corcoran (Eds.), *Measures for clinical practice and research: A sourcebook* (4th ed., Vol. 1, pp. 344–347). New York: Oxford University Press.

Lindof, T. R., & Taylor, B. C. (2011). *Qualitative communication research methods* (3rd ed.). Los Angeles: Sage.

Lunceford, J.K. & Davidian, M. (2004). Stratification and weighting via the propensity score in estimation of causal treatment effects: A comparative study. *Statistics in Medicine, 23*(19), 2937–2960.

Material, Infant, and Early Childhood Home Visiting (2016). MIECHV issue brief on family enrollment and engagement. Washington, DC: Health Resources and Services Administration: Material, Infant, and Early Childhood Home Visiting Technical Assistance Coordinating Center. Retrieved from https://mchb.hrsa.gov/sites/default/files/mchb/MaternalChildHealthInitiatives/HomeVisiting/tafiles/enrollmentandengagement.pdf.

McClain Sampson, M., Duron, J.F., Mauldin, R.L., Kao, D., & Davidson, M. (2017). Postpartum depression, risk factors, and child's home environment among mothers in a home visiting program. *Journal of Family Studies, 26* (5)*,* 1-10.

Michalopoulos C., Duggan, A., Knox, V., Filene, J.H., Lee, H., Snell, E.K., Crowne, S., Lundquist, E., Corso, P.S., & Ingels, J.B., (2013). *Revised design for the mother and infant home visiting program evaluation.* OPRE Report 2013-18. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Michalopoulos, C., Lee, H., Duggan, A., Lundquist, E., Tso, A., Crowne, S., Burrell, S., Somers, J., Filene, J.H., & Knox, V. (2015). *The Mother and Infant Home Visiting Program Evaluation: Early Findings on the Maternal, Infant, and Early Childhood Home Visiting Program.* OPRE Report 2015-11. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

Mihalic, S. (2004). The importance of implementation fidelity. *Emotional and Behavioral Disorders in Youth, 4*(4), 83–105.

Nassar, F., & Wisenbaker, J. (2003). A Monte Carlo study investigating the impact of item parceling on measures of fit in confirmatory factor analysis. *Educational and Psychological Measurement*, *63*(5), 729–757.

Patton, M.Q. (2002). *Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Patton, M.Q. (2008). *Utilization-focused evaluation,* (4th ed.). Thousand Oaks, CA: Sage.

Prinz, R.J. & Moncher, F. J. (1991). Treatment fidelity in outcome studies. *Clinical Psychology Review, 11(3), 247-266.*

Roeser, A. (2009). Relationship quality in early childhood mental consultation efforts: A potential mediators for behavior changes. Unpublished master's thesis, Johns Hopkins University, Bloomberg School of Public Health, Baltimore.

Rosenbaum, P.R. & Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41-55.

Rosenbaum P.R. & Rubin D.B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician, 39*, 33-38.

Samson, T. Iecovich, E. & Shvartzman, P. (2016). Psychometric characteristics of the Hebrew version of the Professional Quality-of-Life Scale. *Journal of Pain and Symptom Management, 52*(4), 575-581.

Segal, L., Opie, R.S., & Dalziel, K. (2012). Theory! The missing link in understanding the performance of neonate/infant home-visiting programs to prevent child maltreatment: A systematic review. *The Millbank Quarterly, 90*(1), 47-106.

Shea, S.E., Goldberg, S., Davies, D., & Weatherston, D.J. (2012). Reflective supervision self-efficacy scale for supervisees. Unpublished measure.

Shea, S.E., Goldberg, S., Davies, D., & Weatherston, D.J. (2016). A community mental health professional development model for the expansion of reflective practice and supervision: Evaluation of a pilot training series for infant mental health professionals. *Infant Mental Health Journal, 37*(6), 653-669.

Shih, T. & Fan X. (2008). Comparing response rates from web and mail surveys: A meta-analysis. *Field Methods, 20*(3), 249-271.

Schroeder, K., Jia, H., & Smaldone, A. (2016). Which Propensity Score Method Best Reduces Confounder Imbalance? An Example From a Retrospective Evaluation of a Childhood Obesity Intervention. *Nursing Research, 65*(6), 465–474.

Sianesi, B. (2004): "An evaluation of the active labour market programmes in Sweden," The *Review of Economics and Statistics, 86*(1), 133–155.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York, NY US: Oxford University Press.

Sommers, M., Zhu, P., Jacob, R., Bloom, H. (2013). The validity and precision of the comparative interrupted time series design and the difference-in-difference design in educational evaluation. MDRC Working Paper on Research Methodology.

Spielberger, J., Gitlow, E., Winje, C., Harden, A., Banman, A, & Dadisman, K. (2013). Building a system of support for evidence-based home visitation programs in Illinois: Findings from year 2 of the Strong Foundations evaluation. Chicago, IL: Chapin Hall at the University of Chicago.

Spielberger, J., Gitlow, E., Winje, C., Harden, & A., Banman, A. (2013). Building a system of support for evidence-based home visitation programs in Illinois: Findings from year 3 of the Strong Foundations evaluation. Chicago, IL: Chapin Hall at the University of Chicago.

Stamm, B.H. (2009). *Professional Quality of Life: Compassion Satisfaction and Fatigue Version 5 (ProQOL)*. www.proqol.org.

Stamm, B.H. (2010). *The Concise ProQOL Manual*. Pocatello, ID: ProQOL.org.

Stevens, J. (1992). *Applied multivariate statistics for the social sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.

Stuart, E.A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science, 25*(1), 1-21.

Totsika, V. & Sylva, K. (2004). The home observation for measurement of the environment revisited. *Child and Adolescent Mental Health, 9*(1), 25-35.

United States Department of Agriculture (2016). *Rural-Urban Continuum Codes.* Retrieved from https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx.

United States Department of Agriculture (2016). *Rural-Urban Continuum Codes.* Retrieved from https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/documentation.

Vittinghoff, E., & McCulloch, C. E. (2007). Relaxing the rule of ten events per variable in logistic and Cox regression. *American Journal of Epidemiology, 165*(6), 710–718.

Watson, C.L., Bailey, A.E., Storm, K.J. (2016). Building capacity in reflective practice: A tiered model of statewide supports for local home-visiting programs. *Infant Mental Health Journal, 37*(6), 640-652.

West, S.G., Finch, J.F., & Curran, P.J. (1995). Structural equation models with non-normal variables: Problems and remedies. In R. Hoyle (Ed.), *Structural Equation Modeling: Concepts, Issues and Applications*, (pp. 56-75). Newbury Park, CA: Sage.

West, B. T., Welch, K. B., & Galecki, A. T. (2007). *Linear mixed models: A practical guide using statistical software.* Boca Raton, FL: Chapman & Hall/CRC.

Wolf, E. J., Harrington, K. M., Clark, S. L., & Miller, M. W. (2013). Sample size requirements for structural equation models: An evaluation of power, bias, and solution propriety. *Educational and Psychological Measurement, 76*(6), 913–934. https://doi.org/10.1177/0013164413495237

What Works Clearinghouse (2014). Procedures and standards handbook version 3.0. Washington, DC: National Center for Education Evaluation and Regional Assistance.