Evaluating Dyadic Interactions as a Mechanism Driving Caregiver and Child Outcomes: Evidence from a Randomized Controlled Trial of the FIND Intervention

by

Andrea M. Imhof

A dissertation accepted and approved in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

in Psychology

Dissertation Committee:

Philip Fisher, Chair

Maureen Zalewski, Core Member

Elizabeth Skowron, Core Member

Nicole Giuliani, Institutional Representative

University of Oregon

Spring 2025

DISSERTATION ABSTRACT

Andrea M. Imhof

Doctor of Philosophy in Psychology

Title: Evaluating Dyadic Interactions as a Mechanism Driving Caregiver and Child Outcomes: Evidence from a Randomized Controlled Trial of the FIND Intervention

Grounded in evidence that responsive caregiving can buffer the negative effects of stress on young children, there have been a growing number of parenting interventions that target responsive caregiving as a key mechanism of change. One challenge in evaluating intervention efficacy has been a relative lack of valid and reliable observational coding tools that can effectively quantify nuanced dyadic interactions and are sensitive to change over time. The current study employs two novel video-coding methods – Conversational Turns (CT) and Follow-Lead-Other (FLO) – to evaluate changes to dyadic interaction variables in a randomized controlled trial of a strengths-based video-feedback intervention (Filming Interactions to Nurture Development – FIND). This study aimed to understand a) how dyadic interaction measures change across the intervention period; b) how dyadic interaction variables were related to other caregiver and child outcome measures; and c) whether CT or FLO variables mediate intervention-related changes in caregiver and child outcomes.

Contrary to hypotheses, dyadic language counts measured with CT did not change significantly across the intervention period. CT variables were also not significantly related to other child language outcomes. Our second observational coding measure yielded more promising results: there was a marginally significant FIND-related increase in FLO Following behaviors across the intervention period, and FLO Following behaviors significantly predicted several measures of caregiver self-efficacy, stress, and executive function. An exploratory

mediation analysis offered promising evidence that FLO Following may partially mediate FIND-related changes in caregiver teaching self-efficacy; results from the current study suggest only marginally significant effects that warrant further follow-up in the context of a larger study with more statistical power.

Ultimately, the current study offers promising evidence that the FIND intervention increases the extent to which caregivers follow their child's lead during interactions, which may have downstream effects on FIND-related changes in caregiver self-efficacy. This study also highlights the future potential of the FLO Coding Tool as an effective measure of responsive caregiving with broader applications across clinical research.

ACKNOWLEDGEMENTS

This dissertation is the culmination of several years of programmatic work evaluating and developing observational video-coding schemes to turn qualitative, dynamic interactions between children and their parents into relevant quantitative data. The nature of this work is collaborative and team-oriented; accordingly, there are many mentors, research assistants, and colleagues who deserve acknowledgment for their role in this work.

First, I thank my advisor and the chair of this dissertation committee, Dr. Philip Fisher, for his steadfast support and career mentorship throughout the span of my graduate training at the University of Oregon. Your work inspired me to pursue a PhD in Clinical Psychology many years ago, and I am honored and humbled to have had the opportunity to learn from you over the years. I also thank the other members of my committee – Dr. Maureen Zalewski, Dr. Elizabeth Skowron, and Dr. Nicole Giuliani – for their mentorship throughout this project. Many of you have been steadfast supports for my clinical and research work across my graduate career and it has been an honor to learn from you across the many projects you've supported. This work would not have been possible without the mentorship of Dr. Shannon Peake, who helped co-create the FLO Coding Tool and provided invaluable feedback and guidance throughout its development process. I also extend sincere gratitude to Dr. Sihong Liu, whose statistical expertise has been an invaluable resource throughout my graduate career. Her mentorship provided much of the backbone for the analytic plan in the current study and I am incredibly grateful for her guidance, patience, and wisdom in all things statistics.

This project truly would not have been possible without the army of research assistants who worked steadily on video-coding projects with me over the past six years. Over 45 undergraduate research assistants were involved in this work; I would especially like to thank

Heather Ralph, Alex, Bui, Alexandra Gonzalez Espitia, Sebastian Ibanez Sanhueza, and Ethan Avila for their tireless work in the development and coding processes for both the CT and FLO measures. I would also like to thank the study team at the University of Denver for their role in data collection and data cleaning for this large-scale RCT of the FIND Intervention – this project would not have been possible without your partnership.

Finally, I would be remiss not to acknowledge the incredible support from my husband, Lars Österberg, whose patience and unwavering belief in me has kept me going through the harder parts of graduate school. To my graduate student colleagues and friends who have spent countless hours in Zoom rooms with me writing and analyzing data – thank you. I also want to acknowledge my two sons – Axel and Felix – for their boundless energy, enthusiasm, and contagious joy, which has kept me grounded and provided the motivation to pursue this line of work. And finally, a note of appreciation for my dog Juno, who has faithfully been at my side (or curled up at my feet) through the entirety of this project. I appreciate all of you more than I can possibly express.

This study was supported by the Administration for Children and Families (Grant No. YR0056, PI: Watamura), the Center on Parenting and Opioids (National Institute on Drug Abuse of the National Institutes of Health, Award No. P50 2173B4) and the Stress Neurobiology and Prevention Lab (PI: Philip Fisher).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Relationships between caregivers and young children hold a lot of power. Caregiver warmth, sensitivity, and responsivity are highly associated with cognitive and socio-emotional development and can serve as important protective factors against other external stressors and environmental threats (Bornstein & Tamis-LeMonda, 1989; Davidov & Grusec, 2006; Hirsh-Pasek & Burchinal, 2006; Landry et al., 2001, 2006; Raby et al., 2015; Tamis-LeMonda et al., 2001). Evidence from interventions in foster and institutionalized care settings suggests that increases in responsive caregiving before the age of three can reverse the physiological symptoms of toxic stress, normalizing electrical brain activity and cortisol patterns disrupted by early stress exposure (Asok et al., 2013; Bernard, Dozier, et al., 2015; Bernard, Hostinar, et al., 2015; Dozier et al., 2008; Fisher et al., 2007; Miller et al., 2011). Other studies have documented the relationship between maternal warmth, maternal sensitivity, and resilience in the face of poverty and other environmental stressors (Farrell et al., 2017; Kim-Cohen et al., 2004; Vernon-Feagans & Bratsch-Hines, 2013). The consistency of this buffering effect across contexts suggests that responsive caregiving may be an important target for interventions designed to reverse the effects of early stress exposure (Bruce et al., 2013; Farrell et al., 2017).

Despite a growing number of parenting interventions that aim to increase caregiver responsiveness, there is a lack of consensus in the literature about the best ways to quantify the complex interactions between caregivers and their children (Gardner, 2000; Zumbach et al., 2022). Responsive caregiving, maternal sensitivity, serve and return interaction, and contingent responsiveness are frequently referenced terms used across the early parenting literature, yet many researchers use widely different methods for coding and evaluating these constructs. Even among coding methods that claim to quantify similar constructs (e.g., “maternal sensitivity),

researchers have found little shared variance among methods that aim to measure the same things (Bohr et al., 2018). It is even more challenging for researchers to identify observational coding tools to evaluate the efficacy of parenting interventions since the chosen methodology needs to be a) aligned with the intervention's theory of change and b) sensitive to change over shorter periods (e.g., pre- and post-intervention time points) to detect effects (Fok & Henry, 2015; Schindler et al., 2017; Zumbach et al., 2022).

This dissertation is the culmination of years of programmatic work to address limitations in the current landscape of observational coding tools to measure responsive parenting. The current study uses two observational coding tools that were developed to evaluate the Filming Interactions to Nurture Development (FIND) intervention, a light-touch, strengths-based intervention that promotes responsive caregiving behaviors among high-stress, high-risk populations. Specifically, the current study uses data from one of the first large-scale randomized controlled trials of FIND to evaluate how the intervention changes interactions between caregivers and their children during play. We also explore how these dyadic interaction variables relate to other caregiver and child outcomes that change across the intervention period, probing possible mediation effects of these two new measures.

Chapter 1 provides an outline of the historical context and situational factors that led to the current study. I first review the history and development of responsive caregiving interventions, situating FIND within the larger landscape of intervention programs that hail from both attachment and social learning approaches. Citing several reviews of existing responsive caregiving interventions, I discuss key qualities that predict intervention program success, highlighting promising key attributes of the FIND intervention. I then introduce two important components of the FIND theory of change - serve and return interactions and following the

child's lead - providing historical context underlying the development of these two constructs within the domain of caregiver responsivity. The two dyadic interaction measures at the core of this study's methodology build off of these key definitions of responsive care, aiming to align closely with FIND intervention targets. Finally, I introduce the conceptual model for the current study, presenting both study aims and key hypotheses in the context of the existing literature on the efficacy of FIND.

History and Development of Responsive Caregiving Interventions

Responsive caregiving has been an important target for interventions for decades; there is a long history of interaction-focused interventions that hail from both attachment and social-learning approaches (for reviews, see Fisher & Skowron, 2017; Kohlhoff et al., 2022). At their core, both attachment and social learning perspectives are based on similar observations about negative outcomes following children's exposure to harsh, neglectful, or inconsistent parenting styles (Mendez et al., 2016; Patterson, 1982; Samdan et al., 2020; Van Aken et al., 2007; Wiggins et al., 2015). Despite different approaches to changing parenting behavior, interventions from each theoretical camp share many underlying features and have generally seen widespread support across multiple randomized controlled trials.

Social Learning Theory and its Impact on Parenting Interventions

Social learning theory emerged in the 1960s following the classic series of "Bobo Doll" experiments that demonstrated how children learn through observation, imitation, and modeling (Bandura, 1965; Bandura et al., 1963). Subsequent work from Patterson et al. at the Oregon Social Learning Center applied these concepts more directly to parenting contexts, observing that low rates of positivity and high rates of harsh discipline led to disruptive behavior problems in children and escalating rates of negative behaviors across family members over time. These

observations led to the development of *coercion theory* (Patterson, 1982; Patterson & Reid, 1984), and the discovery that parental responsiveness is an effective mechanism of change when targeted through intervention (Fisher & Gilliam, 2012; Taylor & Biglan, 1998). Patterson's work forms the foundation of several responsive parenting interventions that aim to reduce coercive parenting practices and increase consistent positive approaches to discipline (e.g., Incredible Years (Gardner & Leijten, 2017), Parent Management Training (Forgatch & Kjøbli, 2016), and Parent-Child Interaction Therapy (Eyberg et al., 2001).

Over the past two decades, there has been widespread empirical support for interventions grounded in social learning theory. At their core, these interventions aim to improve child behavior by improving parent-child relationships. Many of these programs aim to provide caregivers with relevant skills from social learning theory, cognitive behavioral theory, and developmental theory by teaching parents how to model and reinforce behaviors, manage stress, and build warm relationships with their children (Gardner & Leijten, 2017). Several large-scale randomized controlled trials have been conducted investigating the efficacy of PCIT (Lieneman et al., 2017; Skowron et al., 2024; Thomas et al., 2017; Thomas & Zimmer-Gembeck, 2011), Incredible Years (Pidano & Allen, 2015; Webster-Stratton & Reid, 2010), and Parent Management Training (Forgatch & Kjøbli, 2016; Helander et al., 2024; Kazdin, 2005; Michelson et al., 2013; Mouton et al., 2018). Evidence across contexts, cultures, and geographical regions suggests that these interventions consistently improve parenting behaviors, support caregiver well-being, and reduce negative child behaviors (Fisher & Skowron, 2017; Maciel et al., 2023).

Attachment-Based Responsive Caregiving Interventions

Attachment theory is grounded in Bowlby and Ainsworth's research from the 1960s and 70s that highlights how early parent-child relationships impact a child's socio-emotional

development and behavior (Ainsworth, 1963; Ainsworth et al., 1978; Bowlby, 1969; Sutton, 2019). Ainsworth's research was the first to identify definitions of responsive caregiving that included concepts of caregiver sensitivity and contingent responsiveness (Lyons-Ruth, 1996). Her work proposes that toddlers develop an internal working model that considers their parent a "secure base" from which they can explore the world. Children who do not experience consistent sensitive and responsive interactions with caregivers (e.g., through experiences with harsh, abusive, or neglectful parenting) are at risk of developing insecure or avoidant attachment styles, which can lead to downstream challenging behaviors and negative socioemotional outcomes.

Several attachment-based parenting interventions emerged out of this early work that identified links between parenting sensitivity, attachment security, and psychological outcomes (e.g., Child Parent Psychotherapy (Lieberman & Van Horn, 2004; Lieberman et al., 2015), Attachment and Biobehavioral Catchup (Dozier, 2003; Dozier et al., 2008), and Circle of Security (Cooper et al., 2005)). As a whole, these interventions share aims to increase parent sensitivity and responsiveness to child cues, with the ultimate goal of improving caregiver-child attachment patterns. The evidence base supporting the efficacy of attachment-based interventions is robust. Both Child-Parent Psychotherapy (CPP) and Attachment and Bio-behavioral Catchup (ABC) have shown significant positive effects for parents and children across several RCTs (Bernard, Dozier, et al., 2015; Cicchetti et al., 2006; Dozier et al., 2008; Stronach et al., 2013; Toth & Gravener, 2012), including studies that continue to show evidence of impact after long-term follow-up (Guild et al., 2017; Lieberman et al., 2006; Zajac et al., 2020). Attachment-based approaches to intervention differ widely in methodology and approach, including programs that offer a more directive curriculum (e.g., ABC), video-based feedback (e.g., VIPP; Juffer et al.,

2017), and more intensive therapeutic work with caregivers to address intergenerational trauma (e.g., CPP).

Hybrid Approaches: The Rise of Parent Coaching

Although social learning and attachment theories have different theoretical origins and were long considered incompatible due to tensions between behavioral and psychodynamic orientations (Fisher & Skowron, 2017), the past two decades have brought increasing collaboration and hybridization across intervention models. Patterson and Fisher have identified similarities across theoretical camps that emphasize a shared focus on microsocial parent-child interactions as a key intervention target (Patterson & Fisher, 2002). Along these lines, Dozier and colleagues proposed a more direct convergence of theoretical perspectives to increase the impact of interventions by drawing directly from strategies employed by each approach (Dozier et al., 2002). Increasing studies have evaluated the impact of social learning theory interventions using attachment-based outcomes (e.g., Allen et al., 2014; Fisher & Kim, 2007; O'Connor et al., 2013; Patterson et al., 2010) and several intervention models now directly employ a hybridized approach that includes attachment-based sensitivity training and social-learning based approaches to discipline and behavior management (Hutchings et al., 2023; Juffer et al., 2017; Van Zeijl et al., 2006). The past decade has seen a shift towards a broader emphasis on "relational interventions" rather than an explicit delineation between roots within attachment or social learning theory (Fisher & Skowron, 2017; Toth & Gravener, 2012; Valentino, 2017).

In recent years, there has been a rising interest in the potential of video coaching as a tool for intervention. On principle, video-coaching programs utilize several best practices from each theoretical camp, holding parent-child interaction as a central target and streamlining intervention feasibility. Programs that utilize video feedback operate under the same guiding

framework: caregivers are filmed interacting with their children and then allowed to watch and reflect on these filmed interactions with the support of a therapist or coach. Some programs emphasize behaviors, discipline, and limit-setting (e.g., Interaction Guidance: McDonough, 1995), while others target improving caregiver sensitivity and attachment (e.g., Marte Meo: Axberg et al., 2006; Video-Feedback Intervention to Promote Positive Parenting – VIPP: Juffer et al., 2017, 2018). On the whole, video feedback strategies aim to help caregivers increase both confidence and parenting competence through specific and constructive feedback based on their interactions with their children (Balldin et al., 2018; Fukkink, 2008). Reviews of video-feedback intervention programs highlight their overall success in yielding higher levels of maternal sensitivity and improved positive parent-child interactions, although there is little consistency in the methodology used to evaluate these caregiver-child interaction outcomes (Balldin et al., 2018).

Predicting Contributions to Program Success

Research on the development and implementation of responsive caregiving interventions highlights several important factors that contribute to program success. First, meta-analytic methods confirm the efficacy of this focus on responsive caregiving as a target; programs that focus on improving the quality of parent-child interactions have the largest effect sizes for both parent and child outcomes (Wyatt Kaminski et al., 2008). There is also an increasing emphasis on the importance of ***early*** intervention (i.e., birth to three) in buffering the deleterious effects of environmental stress exposure (Blaisdell et al., 2019; Farrell et al., 2017; Flannery et al., 2017; Shonkoff & Phillips, 2000). While many intervention programs still target preschool-aged children, increasing trends in prevention science start relational interventions at birth and sometimes during the perinatal period (Lieberman et al., 2020; Narayan et al., 2021; Waters et

al., 2015). Feasibility of implementation is also paramount. The past decade of translational research has brought an increasing focus on the need for short-term, scalable interventions that are less time-and-resource intensive than their predecessors (Fisher et al., 2016; Schindler et al., 2017).

The FIND Intervention

Building off prior models rooted in both attachment and social learning theory, the Filming Interactions to Nurture Development (FIND) Intervention is a video-coaching intervention designed to support the development of responsive caregiving behaviors among parents living in high-risk, high-stress environments (Fisher et al., 2016). FIND prioritizes intervention feasibility and scalability: it is a light-touch, short-term intervention that can be flexibly adapted to diverse settings (e.g., primary care, group shelters, and individual households). The FIND core curriculum utilizes strengths-based video feedback to encourage increased responsive caregiving behaviors, aiming to increase parent confidence and self-efficacy by showing caregivers positive examples of interactions with their own children. The intervention primarily targets two dimensions of responsive caregiving: back-and-forth ("serve and return") interactions and caregiver sensitivity to child-led interests within play (i.e., following the child's lead). The FIND theory of change predicts that increasing both *serve and return* and *following* behaviors within dyadic interactions will significantly improve both parent and child outcomes.

Serve and Return: Contingency-Based Responsive Caregiving

The idea of "serve and return" interactions was first introduced by researchers at the Harvard Center for the Developing Child in 2004 (National Scientific Council on the Developing Child, 2004; Shonkoff & Bales, 2011). The term refers to the natural back-and-forth exchanges that occur when young children initiate interactions by babbling, gesturing, or vocalizing and their

caregivers respond to them. Shonkoff et al. (2000) developed the serve-and-return ("game of tennis") metaphor to explain the power of environmental influences on early brain development, highlighting evidence that sustained, responsive, and reciprocal interactions between a caregiver and child can have long-term positive effects on child cognitive, language, and brain development (National Scientific Council on the Developing Child, 2004; 2014).

Numerous coding systems have attempted to quantify back-and-forth exchanges across the decades, predating the "serve and return" terminology. Early systems of evaluating parent-child interactions describe a "dance," where parents and children react and respond to each other (Barnard, 1979; Barnard et al., 1983). The concept of *contingent social responsiveness* builds on this concept of a "dance" by characterizing caregiver-child interactions as mutually dependent, with reciprocal contributions from the infant and caregiver (Dunst et al., 1989). Subsequent research has shown that contingent social interactions – and particularly *verbal* back-and-forth interactions – are a critical precursor to language development, early literacy skills, and growth in other cognitive domains (Bornstein et al., 2015; Bornstein & Tamis-LeMonda, 1989, 1997; C. Dunst et al., 1989; C. J. Dunst & Kassow, 2008; Golinkoff et al., 2015; Kochanska et al., 1999; Raver, 1996; Tamis-LeMonda et al., 2001).

Verbal back-and-forth interactions have become an area of particular interest in recent decades with the popularization of research on "conversational turns" (Beiting et al., 2022; Gilkerson et al., 2017; Romeo et al., 2018; Tamis-LeMonda et al., 2001). Numerous studies utilizing day-long audio recordings of parent-child verbal interaction suggest that verbal back and forth ("conversational turns") serve as a critical mechanism driving expressive and receptive language development in the first five years of life (Anderson et al., 2021; Donnelly & Kidd, 2021; Romeo, Leonard, et al., 2018; Romeo, Segaran, et al., 2018). Parallel evidence from

research involving more intensive manual coding of all serve-and-return behaviors (i.e., both gestures and verbal cues) corroborates the idea that back-and-forth contingent interactions are a driving force of cognitive and language growth for young children (Kuchirko et al., 2018; Tamis-LeMonda et al., 2014; Tamis-LeMonda & Bornstein, 2002).

Following the Child's Lead: Sensitivity-Based Responsive Caregiving

While serve-and-return exchanges between caregivers and children comprise one critical component of responsive caregiving, *caregiver sensitivity* is an equally important component of the FIND curriculum. Caregiver sensitivity and warmth, initially described by Ainsworth et al.'s seminal research on attachment in the 1970s, includes caregivers' emotional connectedness, empathetic responsiveness, and attunement to their child. Sensitivity has consistently been linked to the quality of the caregiver-child attachment, which in turn is correlated with positive mental health and child socio-emotional outcomes (Bigelow et al., 2010; Bretherton, 2013; McElwain & Booth-Laforce, 2006; Stams et al., 2002). Early descriptions of sensitivity emphasize both the quality and appropriateness of the caregiver's response to the child's cues (Ainsworth et al., 1974; Shin et al., 2008).

More recently, attachment-based interventions have morphed this idea of sensitivity and responsiveness to child cues into an emphasis on following the child's lead. The Attachment and Biobehavioral Catch-up (ABC) intervention is one of the first to explicitly emphasize caregivers' following their child's lead, citing caregiver *following* as a way to increase caregivers' nurturing behaviors and minimize parental intrusiveness (Dozier & Bernard, 2017). Results from several randomized controlled trials of ABC reveal that this emphasis on following the child's lead is linked to improvements in internalizing and externalizing problems in children, lower levels of negative affect, and increased child language skills (Bernard et al., 2012, 2017; Lind et al., 2014;

Sprang, 2009). Qualitative interviews with parents following the ABC intervention corroborate the value of this focus on *following*, emphasizing increased confidence in their parenting skills and stronger connections with their children (Bagwell-Gray et al., 2023).

Outside of the attachment-focused literature, a parallel thread of research studies by Maccoby and her colleagues (Maccoby, 1999; Maccoby & Martin, 1983; Parpal & Maccoby, 1985) directly evaluated the impact of parental *following* behaviors on parent-child interaction. Their findings suggested that children whose parents followed their lead during play were more likely to be cooperative when asked to clean up (Parpal & Maccoby, 1985); subsequent studies showed that teaching parents to follow their children's lead increased cooperation and decreased "difficult" and non-compliant behaviors (Davidov et al., 2022; Davidov & Grusec, 2006; Kochanska et al., 2013). Several social-learning-theory-based interventions (e.g., PCIT, Incredible Years) also include components of child-directed interaction (CDI), teaching parents to attend to and follow their child's lead as a part of reinforcing child behaviors (Eyberg & Bussing, 2010; Webster-Stratton & Reid, 2010). Davidov et al (2022) theorize that *following* behaviors call on two aspects of caregiver responsiveness by a) increasing sensitivity to their child's cues and b) emphasizing the reciprocal nature of the parent-child relationship. More recently, "child-led play" has also gathered momentum in educational circles, where evidence links decreased externalizing behaviors and positive socio-emotional and language outcomes to child-centered play in early childhood care settings (Bodrova et al., 2000; Wasik & Jacobi-Vessels, 2017).

Scalability and Flexibility: Implementations of FIND Across Cultures and Contexts

At its core, the FIND intervention targets these two domains of responsive caregiving through a simple, scalable methodology that emphasizes flexibility and can easily be modified to

fit diverse contexts (Fisher et al., 2016). To date, FIND has been implemented widely across the United States, Canada, Australia, and the United Kingdom, and is currently being adapted for use in China. Within the U.S., FIND has also been translated and adapted for use with Spanish-speaking caregivers, fathers, group settings (e.g., family shelters), early childcare settings, and pediatric primary care. Across these diverse contexts, research teams have employed similarly flexible, iterative processes of evaluation to assess feasibility and impact. Much of the existing research base highlighting FIND's efficacy stems from networks of community partnerships and larger learning communities like the Frontiers of Innovation (FOI) group at the Harvard Center for the Developing Child and the Early Childhood Precision, Innovation, and Shared Measurement (EC-PRISM) framework based at the University of Oregon.

This rapid-cycle approach to iterative intervention development and adaptation lies at the heart of the FIND intervention. FIND is built on several key principles and a robust theory of change rooted in increasing responsive caregiving; the details of implementation are largely flexible. Schindler et al., 2017 provide a detailed example of a community-centered adaptation in their description of the development of FIND-Fathers (FIND-F). In a process they refer to as "co-creation," Schindler and colleagues engaged in months of community relationship-building and interviews help guide ways to adapt the program to be more effective with Spanish-speaking Latino fathers. The resulting changes (e.g., translation of materials, including male coaches, reducing the number of sessions) increased the accessibility of the intervention without significantly impacting the FIND theory of change. Similar community-based partnerships, interviews, and focus groups have been implemented across other adaptations of FIND, including the current study.

Findings from FIND: A Growing Evidence Base

Preliminary evidence suggests that FIND significantly improves both parent and child outcomes over time. Early trials of the FIND intervention found significant improvements in parental executive function following the intervention period, including both brain and behavior changes associated with increased inhibitory control (Giuliani et al., 2019). Early pilot trials also found evidence that FIND reduces caregiver stress and improves parenting skills, with more significant changes among caregivers with higher exposure to adverse experiences (Schindler et al., 2017). Several other small studies have been conducted across contexts using the EC-PRISM framework to evaluate outcome changes; although these evaluations have not been formally published, many corroborate positive changes for both caregivers and their children.

In the last three years, results from several larger randomized controlled trials of FIND have been published, showing significant increases in caregiver self-efficacy and sense of competence following the 10-week intervention period (Liu, Phu, et al., 2021). A larger follow-up study on FIND-Fathers corroborated several early findings, highlighting lower levels of parenting stress, increased parent encouragement, and increased sense of identity within the fathering role (Schindler et al., 2021). Noting the differential susceptibility of stress-affected children to caregivers' mental health symptoms, these larger trials have also highlighted the efficacy of FIND among high-risk, vulnerable children. Liu, Fisher et al. (2021) found that FIND is most effective in buffering the impact of caregiver psychopathology symptoms in more biologically sensitive children. Subsequent research has also noted significant improvements in child language outcomes (Imhof et al., 2023) and reductions in child externalizing behaviors following FIND (Schindler et al., 2021), suggesting that FIND leads to wide-spread positive brain and behavior changes for both caregivers and their children.

This Study: Understanding Parent-Child Interaction as a Mediator

Despite clear evidence of improved parent and child outcomes following FIND, it is yet unknown how FIND impacts caregiver-child interactions. Several efforts to measure changes in dyadic interactions following FIND have been attempted, each stymied by methodology challenges. Existing standardized protocols for evaluating parent-child interactions have struggled to effectively capture changes in relevant domains of responsive caregiving (i.e., sensitivity, contingent responsiveness). Past attempts at coding FIND-related intervention data have failed in three primary areas: a) failure to achieve adequate inter-rater reliability b) over-emphasizing aspects of responsive caregiving that were not central tenants of the intervention curriculum and c) choice of coding protocols that were not sensitive to change over time (Imhof, 2019). Ultimately, these failures have highlighted a need for improved video-coding protocols that are a) standardized with adequate coder training materials; b) well-tailored to the FIND theory of change; and c) sensitive to change over short timescales.

This project utilizes two novel video-coding approaches to evaluate how dyadic parent-child interactions change across the FIND intervention period. The chosen methodology emphasizes two key aspects of responsive caregiving that are central to the FIND theory of change: a) back and forth ("Serve and Return") interactions and b) the extent to which caregivers follow their children's lead during play. Ultimately, this project aims to understand how dyadic parent-child interactions change across the FIND intervention period, and the mechanistic relationship between these caregiver-child interactions and other parent and child outcomes.

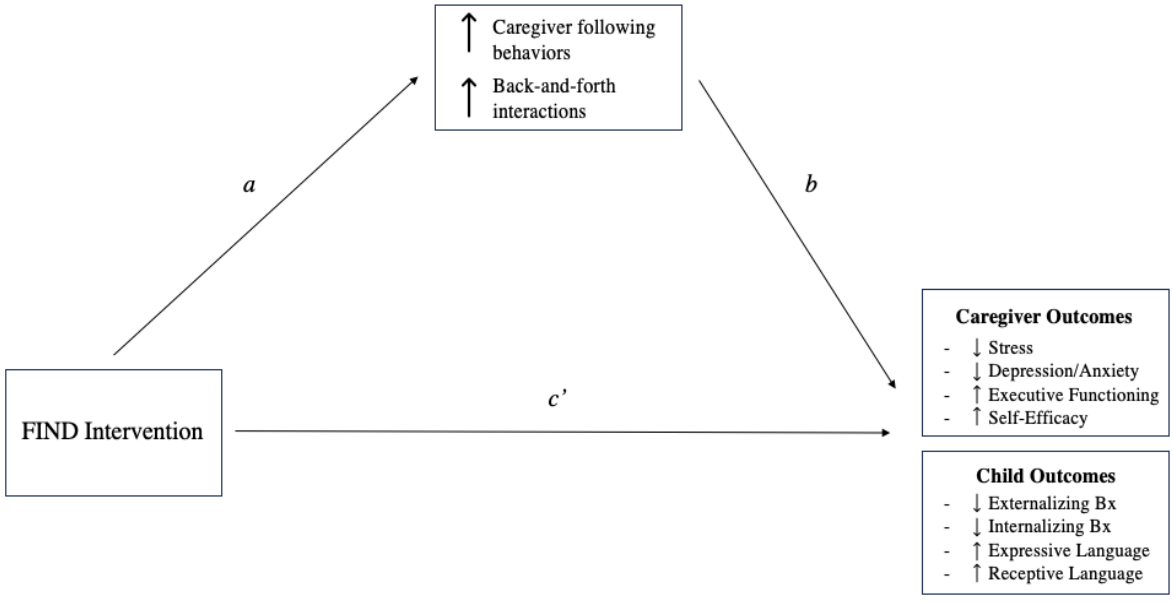
Conceptual Model

The FIND conceptual model places parent-child interaction at the center of its theory of change, honing in on responsive caregiving as a central mediator of caregiver and child

outcomes. This conceptual model and its associated mediation pathways are illustrated in Figure 1. This study aims to clarify each path within the model, first determining the impact of FIND on two domains of responsive caregiving: the quantity of serve-and-return interactions and the extent to which caregivers follow their child's lead during play (*Path A* & *Aim 1).* Aim 2 (*Path B*) examines the relationship between caregiver-child interactions and key caregiver and child outcome measures following FIND. It is expected that parent-child interactions will influence a) post-intervention caregiver outcomes including stress, mental health symptoms, executive function, and self-efficacy; and b) post-intervention child outcomes including language skills and externalizing/internalizing behaviors. Finally, Aim 3 (*Path C')* examines the extent to which FIND-associated changes in parent-child interaction mediate changes to child and caregiver outcomes following the intervention.

Figure 1

Dissertation Conceptual Model



Study Aims & Hypotheses

Aim 1.

Aim 1 evaluates how dyadic interactions change for families in the FIND group compared to the Control group across the intervention period. We expect that dyadic interactions among FIND families will be significantly different from families in the control condition following the intervention period across both measured domains (i.e., back-and-forth interactions and caregiver-following behaviors). Specifically, we expect to see increased Child Utterances (*Hypothesis 1a)*, decreased Parent Utterances (*Hypothesis 1b)*, and increased back-and-forth verbal interaction between caregivers and children (*Hypothesis 1c)* for families in the FIND compared to the control group. We also expect that caregivers in the FIND condition will increase the percentage of time that they follow their child's lead during play (*Hypothesis 1d*) and decrease the amount of time that they spend leading during dyadic interactions (*Hypothesis 1e*) compared to the control group.

Aim 2.

Aim 2 aims to clarify the relationships between dyadic variables and other key caregiver and child outcomes. Based on known existing associations between the current's study's dyadic, child, and caregiver measures (see validation of CT and FLO measures in *Methods*), we expect several significant relationships to emerge between post-intervention outcomes. Specifically, we expect all three measured domains of CT (i.e., Child Utterances, Parent Utterances, Conversational Turns) to predict post-intervention child language scores measured through the PLS-5 and ESCS (*Hypotheses 2a - 2c).* We also expect FLO Following behaviors to be significant predictors of caregiver self-efficacy (*Hypothesis 2d),* caregiver stress (*Hypothesis 2e)*, caregiver mental health (*Hypothesis 2f)*, caregiver executive function (*Hypothesis 2g),* child

internalizing and externalizing behaviors (*Hypothesis 2h)* and child language scores (*Hypothesis 2i).*

Aim 3.

We expect that FIND-related change in dyadic interaction (i.e., CT and FLO outcomes from Aim 1) will mediate observed changes in both caregiver and child outcomes. Aim 3 relies on significant effects from Aim 1 and Aim 2 to evaluate how dyadic interaction may serve as a mechanism of change driving other FIND outcomes. Expected hypotheses mirror those from Aim 2: for the CT measure, we predict that increased Child Utterances, decreased Parent Utterances, and increased Conversational Turns will mediate child language outcomes measured by the PLS-5 and ESCS (*Hypotheses 3a - 3c).* For FLO, we anticipate that increases in Following behaviors will mediate increases in caregiver self-efficacy (*Hypothesis 3d)*, decreases in caregiver stress (*Hypothesis 3e),* improvements in caregiver mental health *(Hypothesis 3f)*, increased performance on executive functioning tasks *(Hypothesis 3g)*, decreased child internalizing and externalizing behaviors *(Hypothesis 3h)*, and increases in child language scores *(Hypothesis 3i).* The hypotheses associated with this project and the larger RCT are pre-registered via the Open Science Framework (Imhof et al., 2019).

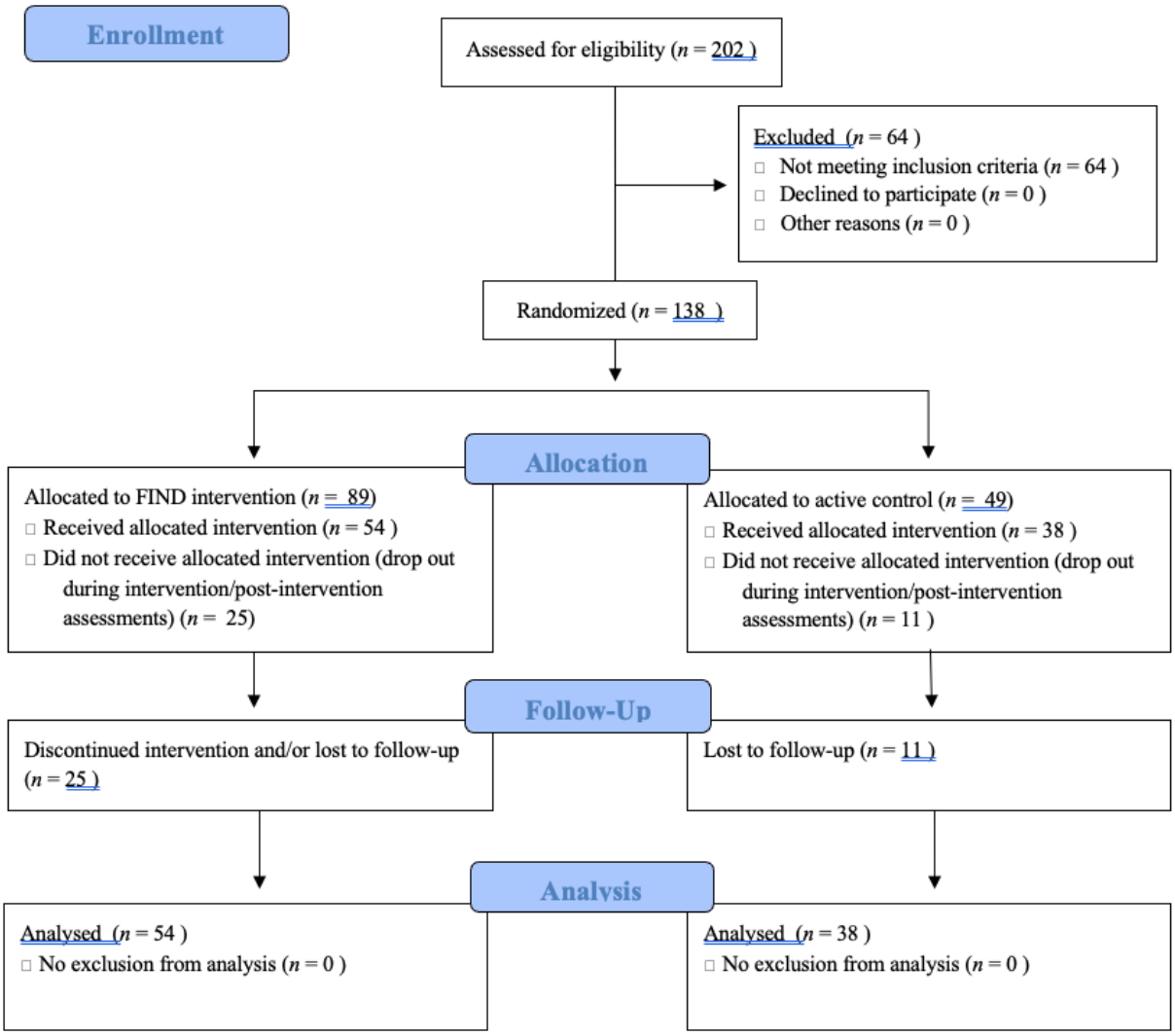
CHAPTER 2: METHODS

Participants, Recruitment, and Screening Methods

Data for this study were collected as a part of a randomized controlled trial (RCT) of the FIND intervention with families recruited from the Early Head Start program in the Denver metropolitan area. 202 families met the initial eligibility criteria, which included a) families with children ages 4-36 months and b) eligibility for Early Head Start services (i.e., low-income families below the federal poverty line). Eligible families then completed a screening visit, which included cortisol assays to determine evidence of cortisol dysregulation among caregivers and children. 138 families were enrolled in the study based on stress criteria (i.e., evidence of cortisol dysregulation) and exclusion criteria (i.e., no diagnosed developmental delays). Enrolled families were randomly assigned into two groups: intervention (FIND + Early Head Start Services, n=89) or active control (Early Head Start services only, n=49). Caregivers and children completed pre- and post-assessment research visits in their homes, which included demographic and behavioral questionnaires, cognitive and behavioral assessments, and filmed dyadic interaction tasks. Of the 138 families enrolled in the study and assigned to a study condition, 92 completed all study visits and were included in this study. Figure 2 outlines the breakdown of participants included in this study's analyses.

Figure 2

Flowchart of Study Participants



Note. 202 participants were screened for eligibility for this study. 138 participants were included and randomized into intervention ("FIND") or active control groups. 92 participants (54 FIND, 38 control) completed the full study protocol including follow-up assessments and are included in this study.

Random Assignment and Blinding Protocols

Following the screening process, 138 eligible participants were assigned to two conditions using conditional randomization, with a greater probability of families being assigned to the FIND intervention condition than the active control group. Because this study included

vulnerable families with evidence of stress-related cortisol dysregulation, conditional randomization was a deliberate decision designed to maximize the number of families who received the active intervention condition. Families assigned to the active control group received standard services from Early Head Start programs without additional support or interventions. Families assigned to the FIND group received the 10-week FIND intervention in addition to Early Head Start services. Throughout study protocols, all researchers except the principal investigator and lead study coordinator were blinded from group assignment and remained unaware of participants' assigned condition to ensure that in-person assessments, data cleaning, and video-coding were unbiased. Recruitment, pre-, and post-intervention assessments for this study occurred between August 15, 2012 and June 10, 2018.

The FIND Intervention

The FIND Intervention is centered around five core elements of responsive caregiving behaviors: a) *Sharing the Focus*, where caregivers are encouraged to notice and respond to the focus of their child's interests; b) *Supporting and Encouraging*, which emphasizes offering help, support, and encouragement in response to child interests; c) *Naming*, where caregivers are taught to offer labels for objects, actions, and feelings in response to the child's interests; d) *Back and Forth*, which encourages caregivers to engage in extended reciprocal, responsive serve and return interactions; and e) *Endings and Beginnings,* where caregivers are taught to notice and respond to their child's changing interests as they end one activity and begin the next. Caregivers enrolled in the FIND intervention program met weekly with a coach, who provided individualized instruction on the FIND core elements across ten weekly home visits. During each home visit, coaches provided strengths-based feedback to caregivers using short, edited video clips that highlight positive examples of the caregiver using the FIND five elements. Coaches

also recorded new examples of caregiver-child interaction during routine activities (e.g., play, meals), which were edited for use during the next session. Sessions were designed to encourage mastery of each FIND core element, build confidence in caregivers' existing skills, and promote increased responsive, serve-and-return interactions between caregivers and their children across the intervention period.

FIND Coaches and Coach Training

The FIND coaches for this research study included five trained therapists or social workers affiliated with the Mental Health Center of Denver or other mental health agencies in the Denver area. All coaches completed a 3-day FIND training that included an overview of the intervention structure, the five FIND elements, and video-coaching technology. Coach training also included opportunities to role-play coaching sessions associated with each FIND element. Throughout the research study, coaches met regularly for group supervision with FIND trainers to continue to build coaching skills, ask and answer questions, and offer increased opportunities to maintain fidelity to the FIND curriculum and intervention protocols.

Intervention Fidelity Coding

To assess intervention implementation fidelity, all coaching sessions with participants were filmed and rated by FIND trainers using a fidelity coding system created for each FIND element. Fidelity coders were asked to evaluate four key aspects of each coaching session: a) review of previous FIND elements, b) discussion of childcare notes, c) introduction of a new element, and d) frame-by-frame analysis of an edited film. Fidelity coders marked "yes" or "no" to several key activities in each session, and each of the four main activities in a session received an overall score (1=yes, 0=no). Averages of the four activity scores were calculated to create a

score for each individual coaching session, and an overall average fidelity score was created for each family.

To ensure coding reliability across fidelity coders, six fidelity coders engaged in extensive training, rating and coming to consensus on all session films for two families - one in English and one in Spanish. Coders were assigned to code independently after 4 weeks of training, when inter-rater reliability reached 75% percent agreement or higher. 20% of all coaching session films were double-coded to evaluate inter-rater reliability for the intervention fidelity coding process. Results from fidelity coding confirmed that families who received the FIND intervention received a high intervention dosage (M=4.69 out of five elements, SD = 0.61), with the majority of families (n = 40, 74.1%) completing intervention sessions for all five FIND elements. Fidelity to the intervention was also high: for all sessions that were coded for fidelity, 58.7% of sessions reached 100% fidelity and 83.0% of sessions reached at least 75% fidelity. Inter-rater reliability for all double-coded fidelity sessions was also high (89% agreement).

Research Study Visits

Families completed three research study visits in their homes at screening, pre-intervention, and post-intervention time points; families were compensated $50 for each research visit. During the screening visit, caregivers completed several study questionnaires about household characteristics, family demographics, exposure to adverse events, and their child's developmental history. Caregivers and children also provided salivary cortisol samples that were used to determine eligibility for enrollment in the study.

After enrollment, families completed two additional research visits in their homes at pre- and post-intervention time points. During these visits, caregivers completed several additional

questionnaires about their own experiences and their child's behaviors, development, and developmental history. Caregivers and children were also asked to complete individual behavioral assessments (e.g., Wisconsin Card Sort Task, Preschool Language Scales). Finally, caregiver-child dyads were filmed completing several interactive activities together (e.g., book reading, free play, and a series of researcher-controlled stressor tasks). Salivary cortisol samples were taken in between each dyadic task within the overall stress paradigm; cortisol samples are not included in the analysis for this specific study.

Home Language & Language of Administration

One noteworthy aspect of the current study is the demographic makeup of participating families: the majority of families participating in this study (n = 74, 80.4%) identified as Hispanic/Latinx and reported speaking some Spanish at home. Caregivers reported what language they preferred to speak at home during the initial screening visit on a 5-point Likert scale that ranged from "Spanish all the time" to "English all the time." All relevant measures associated with this study were available to participants in either English or Spanish; families who indicated Spanish as their preferred language (n = 63, 68.5%) conducted their research visits and intervention sessions in Spanish, with bilingual and bicultural coaches and research personnel.

Measures

This randomized-controlled trial of the FIND intervention collected hundreds of measures to capture relevant information about each caregiver, child, and their dyadic functioning. These measures were collected during two research visits to the family's home that took place before and after the 10-week intervention period. Measures relevant to this project's specific aims are discussed in detail below; note that other biological measures (e.g., salivary

cortisol) were also collected during these visits but are not included in this project. In addition to the instruments outlined below, caregivers also completed a series of demographic questionnaires that documented age, sex, ethnicity, languages spoken at home, and information about income, employment history, and relevant measures of economic hardship, food security, physical health, and housing.

Caregiver Self-Efficacy

Caregivers' self-efficacy was measured using two questionnaire-based measures: a) the Parental Sense of Competence Scale (PSOC; Johnston & Mash, 1989) and b) the Self-Efficacy for Parenting Tasks Index - Toddler Scale (SEPTI-TS; Coleman & Karraker, 2003). The PSOC is a 17-item questionnaire that offers statements relevant to caregivers' confidence in their parenting abilities (e.g., "*If anyone can find the answer to what is troubling my child, I am the one*") and satisfaction with their role as parents (e.g., "*Even though being a parent could be rewarding, I am frustrated now while my child is at his/her present age").* Respondents rate how much they agree with each statement on a scale of 1 ("*Strongly Disagree"*) to 6 ("*Strongly Agree"*). Several items on this scale were reverse coded and a total score was created based on the sum of all scores, with higher scores indicating a higher sense of parenting competency. Within the PSOC, two subscales highlight two domains of parenting competency: the Satisfaction Scale and the Efficacy Scale. This study used the combined total score from both scales as an estimate of overall parenting satisfaction and self-efficacy. The value for Cronbach's Alpha for the PSOC was $\alpha = 0.83$ at pre-intervention and $\alpha = 0.82$ at post-intervention.

The SEPTI-TS is a 53-item questionnaire that assesses seven domains of parent self-efficacy: a) emotional availability; b) nurturance & empathetic responsiveness; c) protection

from harm; d) discipline and limit setting; e) play; f) teaching; and g) instrumental care and establishment of structure/routines. This study used an abbreviated 32-item version that focused on four of these domains most relevant to the FIND intervention: a) nurturance and empathetic responsiveness ("*My child knows that I understand when his/her feelings are hurt"*) - 8 items; $\alpha_{pre} = 0.61$, $\alpha_{post} = 0.62$; b) discipline and limit setting ("*Setting limits for my toddler is relatively easy for me"*) – 7 items; $\alpha_{pre} = 0.72$, $\alpha_{post} = 0.78$; c) teaching ("*I believe that my child learns a lot while I show her/him examples and teach her/him things"*) – 9 items; $\alpha_{pre} = 0.59$, $\alpha_{post} = 0.67$; and d) establishment of structure and routines ("*I have worked out a fairly regular morning routine with my toddler"*) – 8 items; $\alpha_{pre} = 0.76$, $\alpha_{post} = 0.77$. Scores were summed to create a total score and four domain scores, with higher numbers indicating higher parental self-efficacy. The SEPTI-TS has been validated across diverse populations and is now widely used across clinical and research settings (van Rijen et al., 2014; Vance & Brandon, 2017).

Caregiver Stress

Caregiver stress was measured using the Parenting Stress Index - Short Form, 4th edition (PSI-SF-4; Abidin et al., 2006), a 36-item questionnaire that quantifies three major subscales of stress: a) Parental Distress (PD), which measures the extent to which parents feel competent, supported, or conflicted about their role as a parent ($\alpha_{pre} = 0.80$; $\alpha_{post} = 0.92$); b) Parent-Child Dysfunctional Interaction (P-CDI), which measures parent satisfaction with their child and their interactions with them ($\alpha_{pre} = 0.84$; $\alpha_{post} = 0.87$); and c) Difficult Child (DC), which considers whether the child is easy or difficult to take care of ($\alpha_{pre} = 0.86$; $\alpha_{post} = 0.88$). A Total Stress score was also calculated, which indicates the overall level of stress the caregiver was feeling relative to their role as a parent. Caregivers responded to items on a five-point scale that ranged from SA ("*Strongly Agree"*) to SD (*"Strongly Disagree"*). Items were added to create each raw

subscale score, with higher scores suggesting higher levels of stress. Percentile scores were also calculated to provide a standardized comparison to a norm-referenced sample. The measure is widely used in clinical and research settings, with demonstrated reliability and validity across diverse populations (Ríos et al., 2022).

Parents also completed the Parent Daily Report (PDR; Chamberlain & Reid, 1987), a 34-item questionnaire that measures both the frequency of children's challenging behaviors and the extent to which those behaviors are perceived as stressful. Caregivers are asked if specific behaviors have occurred in the past 24 hours (e.g., "*defiance," "whining," "hitting others"*). For each item, caregivers marked whether the behavior occurred and whether or not they considered that behavior problematic or stressful. The internal consistency of this measure was adequate ($\alpha_{pre} = 0.86$, $\alpha_{post} = 0.67$). Caregivers also completed additional questions that asked them to rate their overall experiences of stress during the 24-hour reporting period, their experiences of their child's behaviors as hassles, and an overall rating of positive and negative mood across the previous 24-hour period. For this study, these four questions (Daily Stress, Daily Hassle, Positive Mood, and Negative Mood) were used as a proxy for caregiver stress in response to their child's daily behaviors.

Caregiver Mental Health

Caregivers' symptoms of depression were assessed using the Center for Epidemiological Studies Depression Scale (CES-D; Radloff, 1977), a widely-used questionnaire that has been validated across ages and diverse ethnic and cultural groups (Lewinsohn et al., 1997; Roth et al., 2008). The 20-item scale asks respondents to rate how often they experienced depressive symptoms in the last week and is based on the primary symptom clusters of Major Depressive Disorder: sadness (dysphoria), loss of interest (anhedonia), appetite, sleep,

thinking/concentration, guilt (worthlessness), fatigue, movement, and suicidal ideation. Respondents responded to statements (e.g.. "*I felt that I could not shake off the blues even with help from my family or friends*") using a scale that ranged from 0 ("*Rarely or none of the time*") to 3 ("*Most or all of the time*"); internal consistency for respondents within this sample was adequate ($\alpha_{pre} = 0.61$, $\alpha_{post} = 0.89$). A total score was calculated that ranged from 0 to 60, with higher scores suggesting higher rates of depression. A cutoff score of 16 or higher was used to suggest clinically significant levels of depression.

Caregivers' symptoms of anxiety were assessed using the General Anxiety Disorder - 7 (GAD-7; Spitzer et al., 2006), a brief seven-item measure frequently used to characterize the severity and range of anxiety symptoms. The GAD-7 is widely used in clinical and research work and has been validated for use with diverse populations (Lowe et al., 2017). The GAD-7 asks respondents to rate how often they have experienced seven common anxiety symptoms in the prior two weeks (e.g., "*Feeling nervous, anxious, or on edge*") from 0 ("*Not at all*") to 3 ("*Nearly every day*"); internal consistency for respondents within this sample was high ($\alpha_{pre} = 0.84$, $\alpha_{post} = 0.85$). A total score was calculated that ranges from 0 to 21, with higher scores suggesting higher rates of anxiety. A cutoff score of 10 was used to suggest clinically significant levels of anxiety.

Caregiver Executive Function

In addition to questionnaires, caregivers completed two behavioral tasks to evaluate executive functioning and inhibitory control. The Wisconsin Card Sort task (Grant & Berg, 1948) is a standardized neuropsychological test that measures executive function in terms of concept formation, planning, and cognitive flexibility. This study utilized a computerized version of this task (Heaton, 1993), which caregivers completed during pre- and post-intervention study

visits on researcher-provided laptops. Computerized versions of this task have been found to be psychometrically equivalent to the original version (Artiola i Fortuny & Heaton, 1996; Feldstein et al., 1999). Caregivers were presented with a series of four numbered stimulus cards with patterns that varied by color, shape, and number. They were told to sort the numbered response cards according to different undisclosed sorting rules; respondents were not given explicit instructions about how to sort of the cards, only corrective feedback about whether they had sorted the cards correctly. Throughout the task, sorting rules would change, and caregivers were expected to flexibly adapt their approach to sorting throughout the test administration. Responses were tallied in four categories: a) correct; b) errors; c) perseverative errors; d) non-perseverative errors. Perseverative errors were defined as responses to newly shifted categories that would have been correct for the immediately preceding category. Total raw and standardized scores were automatically calculated as part of the computerized WCST. Broadly, higher scores on the WCST are considered to reflect stronger abstract reasoning abilities and strengths in several related domains of executive functioning including set-shifting, strategic planning, inhibition of impulsive responding, and goal-oriented behavioral responding (Chelune & Baer, 1986). This study used four scores to evaluate performance across the WCST: the total number of trials completed, overall percent accuracy, the percentage of perseverative responses, and the percentage of responses aligned with conceptual categories.

The second computerized behavioral task that caregivers were asked to complete was the N-back task (Owen et al., 2005), a continuous performance task that measures working memory capacity. Caregivers were presented with a sequence of letters on the screen and asked to press a button when the current letter matched the one from *n* steps earlier in the sequence. For example, during the 1-back sequence, caregivers were asked to identify when the letter immediately

beforehand was the same as the current letter on the screen ("one back"); during the 2-back sequence caregivers identified when the letter presented "two back" matched the current letter. Caregivers were given four conditions: zero-back, one-back, two-back, and three-back. Percent accuracy, number of errors, and reaction times were collected across conditions; fewer errors, higher accuracy and faster reaction times generally indicate higher levels of working memory capacity and executive functioning abilities. This study used four measures as a proxy for performance across working memory contexts: a) percent accuracy on the 2-back task; b) percent accuracy on the 3-back task; c) overall percent accuracy; and d) the ratio of hits to false alarms.

Child Behaviors & Socio-Emotional Outcomes

Caregivers also completed a series of questionnaires that reported on several domains relevant to their child's behavior and socio-emotional functioning. The Parent Daily Report (PDR) Total Behavior Score (described in detail in *Caregiver Stress*) offered insight into the number of problematic behaviors that the child engaged in over the past 24 hours. The Total Behaviors score – a tally of all endorsed child behaviors – was used as a proxy for child externalizing behaviors in the current study($\alpha_{pre} = 0.85$, $\alpha_{post} = 0.86$).

Other child behavior outcomes were measured using a composite of two other scales: the Child Behavior Checklist - Preschool version (CBCL/1.5-5; Achenbach, 2001) and the Brief Infant-Toddler Social Emotional Assessment (BITSEA; Briggs-Gowan & Carter, 2006). The CBCL/1.5-5 is a 99-item checklist that documents child internalizing and externalizing behaviors. Caregivers were asked to describe how often their child engaged in specific behaviors on a scale of 0 ("*Not True")* to 2 (*"Very True or Often True")* within the past 2 months. The checklist included a variety of items across externalizing (e.g., "*Physically attacks people*"; $\alpha_{pre} = 0.82$, $\alpha_{post} = 0.78$) and internalizing (e.g., "*Nausea, feels sick (without medical cause*"; $\alpha_{pre} =$

0.74, α_{post} = 0.64) behaviors. The BITSEA is a 42-item questionnaire that is used to evaluate socio-emotional and behavioral problems in clinical and research settings. Like the CBCL, the BITSEA questions ask caregivers to describe their child's behaviors over the previous month on a 3-point Likert scale from 0 ("*Not True"* to 2 ("*Very True")*; questions assess the frequency of externalizing problems (e.g., aggression, defiance; α_{pre} = 0.68, α_{post} = 0.61) and internalizing problems (e.g., anxiety, depression; α_{pre} = 0.64, α_{post} = 0.46) along with dysregulation (e.g., eating and sleeping problems), maladaptive behaviors, and atypical behaviors.

Due to differences in age norms across these two measures, the CBCL was given to children 18 months and older (pre: 65.93%; post: 89.87%) and the BITSEA was administered to families with children under 18 months (pre: 34.07%; post: 10.13%). A composite score was calculated that reflected the percentage of endorsed items out of the total possible score across the CBCL and BITSEA internalizing and externalizing domains, adjusting for missing items. The composite scores for internalizing and externalizing behaviors were used in analyses for this study.

Child Language

In addition to questionnaire-based measures, children completed two behavioral assessments of language competency: the Preschool Language Scales-5 (PLS-5; Zimmerman et al., 2011) and the Early Social-Communication Scales (ESCS; Mundy et al., 2003). The PLS-5 is a norm-referenced and standardized language assessment for children ages 0 to 7 with items that range from pre-verbal interaction-based skills to emerging language and early literacy. It consists of two primary subscales: a) Auditory Comprehension, which quantifies children's abilities to understand gestures, vocabulary, and language concepts; and b) Expressive Communication, which captures children's ability to produce gestures and language to express themselves and

describe objects and actions. Scores from both domains were summed and three standardized scores were calculated (i.e., Total Language, Expressive Communication, Auditory Comprehension) to provide an estimate of expressive, receptive, and overall language abilities. For families who reported Spanish as their primary home language, the PLS-5 Spanish (PLS-5S; n=63; 68.5%) was administered. In this version of the PLS-5, all items are administered first in Spanish; all items that receive a score of zero are re-administered in English, allowing the opportunity for children to hear questions and respond in either English or Spanish. This method – which differs from the monolingual model utilized with other measures collected in this study – offers a more complete representation of a bilingual child's language mastery than single-language scoring (Peña et al., 2016). The internal consistency of the PLS, as measured by Cronbach's alpha, was acceptable across the intervention period ($\alpha_{pre} = 0.61$, $\alpha_{post} = 0.79$).

Children also completed the Early Social-Communication Scales (ESCS), a videotaped and structured observation that assesses verbal and nonverbal communication skills in children between 8 and 30 months. The ESCS protocol dictates a standard set of toys designed to elicit social interactions, joint attention, and behavioral requests: wind-up mechanical toys, cars and balls that move, books, colorful posters, and a toy comb, hat, and glasses. Children were presented with 25 semi-structured situations that encouraged interaction between the researchers and the child. Videotaped records were subsequently coded and scored based on developmental stage, communicative goal, and whether the child initiated the interaction or responded to the researcher's bid. Three categories of children's behaviors were scored: a) joint attention behaviors (i.e., sharing the focus with a researcher); b) behavioral requests (i.e., using verbal and nonverbal behaviors to obtain an object or reach a goal c) social interaction behaviors (i.e., engaging in playful turn-taking interactions with others). Across these three categories, behaviors

are scored based on whether they are child-initiated (i.e., Initiating Joint Attention ($\alpha_{pre} = 0.79$, $\alpha_{post} = 0.81$), Initiating Behavioral Requests ($\alpha_{pre} = 0.77$, $\alpha_{post} = 0.74$), Initiating Social Interaction ($\alpha_{pre} = 0.71$, $\alpha_{post} = 0.69$), or child responses to the researcher's communicative bid (i.e., Responding to Joint Attention ($\alpha_{pre} = 0.21$, $\alpha_{post} = 0.11$), Responding to Behavioral Requests ($\alpha_{pre} = 0.60$, $\alpha_{post} = 0.65$), Responding to Social Interaction ($\alpha_{pre} = 0.60$, $\alpha_{post} = 0.56$). As with the PLS-5S, children who spoke Spanish at home were offered the opportunity to interact with the researcher in both English and Spanish, offering a more complete picture of bilingual communication mastery than a monolingual presentation alone.

Dyadic Free-play Interactions

Caregivers were asked to complete a series of filmed interaction tasks with their child, using a standardized set of age-appropriate toys provided by the researcher. This project utilized 7-minute film clips of caregivers playing with their children and the provided bag of toys. The range of toys offered as options within the provided bags included: a barn with animals, a musical airplane, and a set of play utensils and food of different shapes and sizes. This task took place in the middle of a series of other filmed dyadic interaction tasks, immediately after book-reading. Caregivers were instructed to play with the provided toys as they normally would play with their child at home.

Observational Coding Measures for Dyadic Free-play Interactions

This study utilized two novel observational coding protocols that were specifically developed to quantify behaviors within dyadic play interactions relevant to the FIND theory of change. Each coding protocol was created via an iterative development process that included extensive pilot testing, conferencing, and collaborative discussion across diverse film contexts. These standardized coding protocols were created, tested, and validated prior to their application

within this study. Teams of undergraduate research assistants were trained to reliability on a series of training videos developed for each coding protocol. Training for the Conversational Turns measure took approximately 2 weeks and required coders to reach a 90% match with one training video. Training for the FLO measures took place across a 4-week period using five Gold Standard training videos. Coders needed to reach at least 80% agreement on three training videos before they could begin coding independently. 20% of all films were double-coded and checked to ensure adequate inter-rater reliability (>80% agreement and *kappa* > 0.7).

Conversational Turns (CT).

The Conversational Turns (CT) coding tool was developed as a simple and effective way to quantify utterances and vocalized serve-and-return interactions between a caregiver and child. The protocol was designed for coding use within ELAN Linguistic Annotator software, a program specifically designed for language-based coding of audio and video recordings. It was intended to be a simple, straightforward coding scheme that could be applied across a variety of contexts to quantify caregiver and child language. Coders set up two language tiers in ELAN for utterance coding: one to capture utterances made by the target child and one for the caregiver. Segments of film that contained more than two people (e.g., siblings, researcher interactions) were excluded from analysis. Figure 3 includes the CT coding rules that served as guidance for what counts as an utterance, where utterances start and stop, and for whom utterances should be coded. Reliability across all coded films (n = 29) was very high; inter-rater reliability was measured three ways: a) percent match between coders at the 0.5-second level; b) Cohen's kappa; and c) Pearson's r calculated by aggregating utterance data into 15-second bins. Reliability was calculated separately for Parent Utterances (91.75% match; r = 0.898, κ = 0.805) and Child Utterances (95.73% match; r = 0.859, κ = 0.719).

Figure 3

Coding Rules for the CT Coding Tool

a) Mark any sound that you think was made by the vocal tract (i.e., larynx, throat, mouth, lips, etc.) In order to count as an utterance, the sound must have communicative intent (e.g., words, sound effects, gurgling). Coughing, sneezing, and involuntary sounds from the vocal tract should not be marked as utterances.

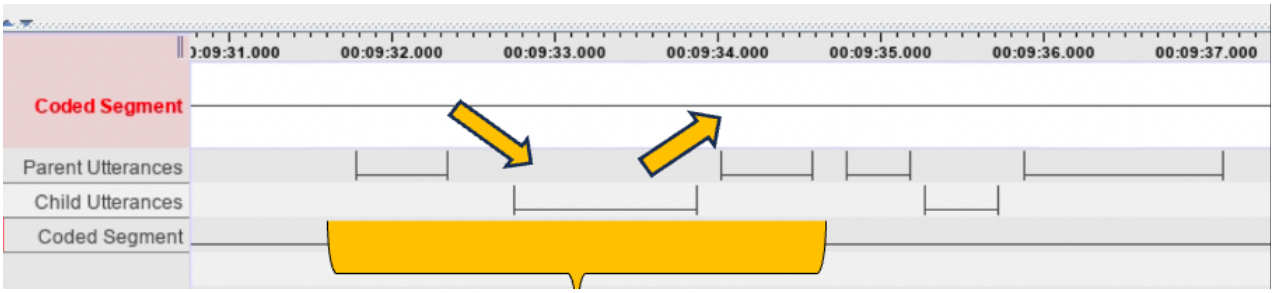
b) Be as accurate as possible when setting utterance boundaries – try to be within tenths of a second. Pauses between phrases that are greater than 0.25 seconds and/or obvious changes in sound effects to indicate a change in concept should be marked as separate utterances.

c) Caregiver and child utterances should ONLY be coded for the target child and target caregiver in the dyadic interaction. Do not mark utterances from other adults, children, or electronics (including toys, radio, music, or TV).

After behavioral coding was completed, coded data was exported and run through the chattr package in R (Casillas & Scaff, 2021). Six primary metrics were calculated: Number of Child Utterances, Mean Length of Child Utterances, Number of Adult Utterances, Mean Length of Adult Utterances, Total Conversational Turns, and Conversational Turns per Minute. "Conversational Turns" were defined as a three-utterance sequence where child and adult speakers took turns engaging in verbal back-and-forth interaction (i.e., parent-child-parent or child-parent-child sequences of utterances). Figure 4 shows an example of a parent-child-parent conversational turn (in orange) taken as a screenshot from the ELAN coding interface. Turns were calculated using a latency threshold of 5 seconds to establish contingency between speakers, consistent with the definition of conversational turns employed by LENA and the established precedent by other researchers who have published on this construct (Donnelly & Kidd, 2021; Ferjan Ramírez et al., 2021; Romeo, Leonard, et al., 2018).

Figure 4

Manual Coding of Utterances and Conversational Turns



Prior Validation of the CT Coding Tool.

Following the development of the CT protocol, some preliminary validation work was completed that offers insight into the relationship between parent-child utterances, verbal back-and-forth, and other parent and child outcome measures. Correlations between the six CT metrics and related caregiver outcomes are reported in Table 1. Most notably, Child Utterances are significantly correlated with several subscales of the Parenting Stress Index: Dysfunctional Interaction, Difficult Child, and Total Stress, suggesting that higher levels of parental stress are significantly correlated with increased Child Utterances. Parent Utterances are also significantly correlated with the PSI Dysfunctional Interaction and Difficult Child scales, while Conversational Turns are significantly related to the PSI Difficult Child and Total Stress subscales. Overall, there is a robust connection between parent stress and increased back-and-forth verbal interaction, although causality and directional effects remain unclear. Parent Utterances are negatively correlated with discipline and limit setting, suggesting that parents who feel most confident about their ability to set limits also tend to talk less during play interactions. Finally, it should be noted that there is a strong correlation between caregiver self-reported ethnicity and observed Parent Utterances during play, with Hispanic caregivers speaking significantly more than their non-Hispanic counterparts during dyadic play.

Table 1

Correlations between Caregiver Outcome Measures and Conversational Turns

Caregiver Measures	Subscales	# Child Utterances	*M* Dur Child Utterances	# Parent Utterances	*M* Dur. Parent Utterances	CT	CT / min
Stress	PSI Parental Distress	0.070	-0.026	0.071	0.076	0.073	0.077
	PSI Dysfunctional Interaction	0.179*	0.027	0.174*	0.045	0.156	0.154
	PSI Difficult Child	0.301**	0.067	0.179*	0.092	0.303**	0.288**
	PSI Total Stress	0.228**	0.029	0.169	0.090	0.224*	0.218*
	Severity of Economic Stress	-0.138	0.093	-0.128	0.118	-0.141	-0.131
Self-Efficacy	SEPTI Nurturance	-0.257	-0.221	0.049	-0.075	-0.217	-0.224
	SEPTI Discipline/ Limit Setting	-0.020	-0.247	-.376*	0.048	0.14	0.028
	SEPTI Teaching	0.001	-0.114	0.216	0.195	-0.42	-0.045
	SEPTI Structure/Routines	-0.004	-0.285	0.211	0.084	0.064	0.069
	SEPTI Total Score	-0.072	-0.296	-0.301	0.070	-0.039	-0.035
Sense of Competence	PSOC	-0.172	-0.059	0.035	-0.071	-0.127	-0.126
Parent Adverse Experiences	Parent ACES	-0.023	0.081	-0.067	0.133	-0.025	-0.028
Caregiver Ethnicity	Hispanic	0.050	-0.102	0.308**	-0.061	-0.153	0.149

* $p < 0.05$, ** $p < 0.001$

Correlations between CT subscales and related child outcomes are reported in Table 2. Most notably, there are robust associations between Child Utterances, Conversational Turns, and child language outcomes: Child Utterances are significantly correlated with PLS measures of receptive and expressive language, along with the ESCS Initiating Social Interaction subscale, Responding to Joint Attention, and Responding to Joint Interaction. Conversational Turns are significantly correlated with PLS Auditory Comprehension, ESCS Responding to Joint

Attention, and ESCS Responding to Social Interaction. There were also significant associations between Child Utterances, Conversational Turns, and BITSEA externalizing behaviors. Although these associations were not significant across other measures of child externalizing (e.g., CBCL, PDR), the relationship between difficult child behaviors and child verbalizations parallels the associations seen between Child Utterances and the PSI Difficult Child and Dysfunctional Interaction subscales (see Table 1).

Table 2

Correlations between Child Measures and Conversational Turns

Child Measures	Subscales	# Child Utterances	*M* Dur. Child Utterances	# Parent Utterances	*M* Dur. Parent Utterances	CT	CT / min
Internalizing Behaviors	BITSEA Internalizing	0.155	-0.245	-0.24	-0.161	0.221	0.240
	CBCL Internalizing	0.027	-0.015	-0.067	-0.121	-0.007	-0.019
Externalizing Behaviors	BITSEA Externalizing	0.428**	-0.193	0.161	0.139	0.420**	0.381*
	CBCL Externalizing	0.023	0.133	-0.164	0.015	-0.051	-0.058
	PDR Total Problems	-0.019	-0.005	0.066	-0.054	0.004	-0.03
Child Language	PLS Auditory Comprehension	0.291**	0.065	0.165	0.002	0.284**	0.271**
	PLS Expressive Communication	0.226*	0.018	0.109	0.108	0.202	0.191
	ESCS Initiating Behavior Request	0.032	-0.011	0.106	0.178	0.061	0.081
	ESCS Initiating Joint Attention	0.167	0.060	0.031	0.060	0.182	0.191
	ESCS Initiating Social Interaction	0.214*	-0.031	-0.119	0.010	0.158	0.169
	ESCS Responding to Behavior Request	0.131	-0.108	0.086	-0.106	0.174	0.175
	ESCS Responding to Joint Attention	0.435**	0.124	0.071	0.119	0.419**	0.428**
	ESCS Responding to Social Interaction	0.246*	0.065	0.119	0.066	0.229*	0.223*

* $p < 0.05$, ** $p < 0.001$

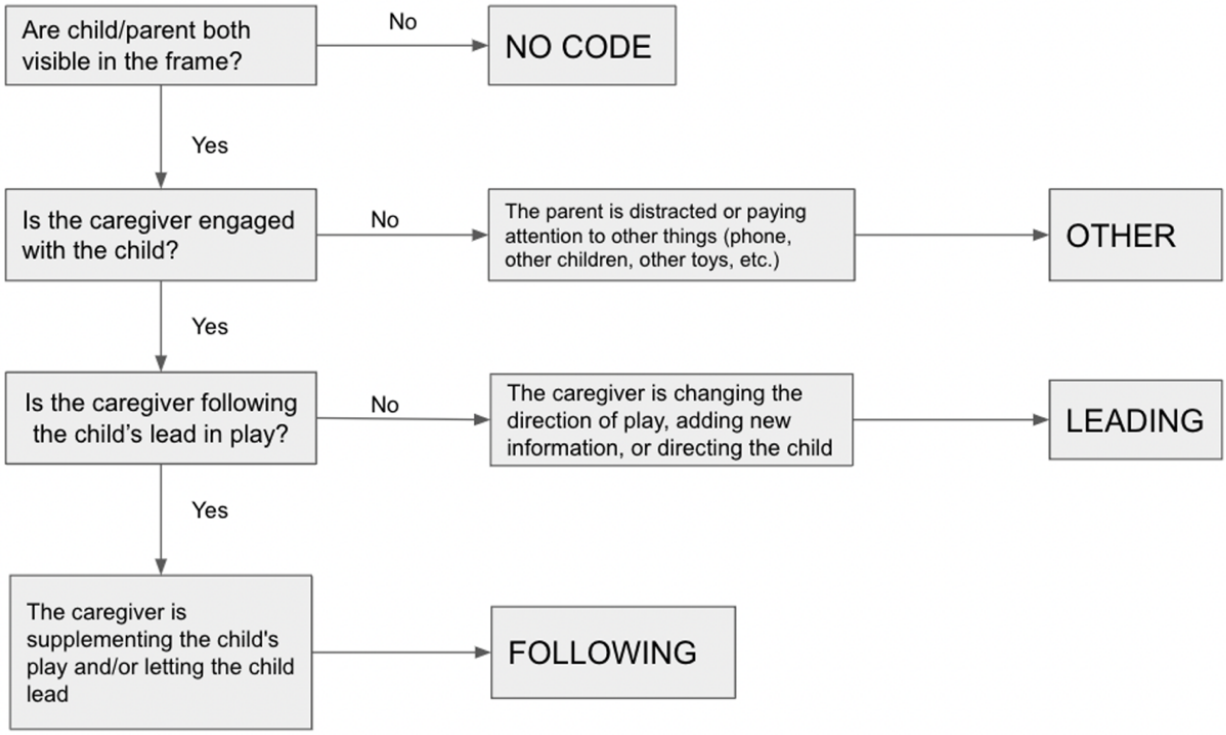
Ultimately, results from this early validation of the CT measure suggest strong associations between Child Utterances and Conversational Turns during interactive dyadic play and other outcomes. Children who verbalize more during dyadic play tend to have stronger auditory comprehension and expressive language skills, initiate social interactions more, and more consistently respond to bids from caregivers. These "talkative" children are also more likely to engage in externalizing behaviors or be rated by their caregivers as "difficult." There is a significant positive association between child verbalization, dyadic conversation, and caregiver stress, and a negative correlation between caregivers' perceived confidence in their ability to set limits and the amount that they talk during play interactions. Based on these associations between Conversational Turns and other relevant measures for this study, three CT subscales were chosen for evaluation within this study: Child Utterances, Parent Utterances, and Conversational Turns.

Follow-Lead-Other (FLO).

The FLO Coding Tool was designed to capture the extent to which caregivers engage in Leading and Following behaviors during dyadic interactions. The protocol was developed for use with Noldus Observer XT, a software program specifically designed for research-based video coding. Figure 5 outlines the FLO decisional flow chart, which explains how the four codes (No Code, Other, Leading, and Following) are utilized.

Figure 5

Decision Flow-Chart Outlining Coding Decisions for the FLO Coding Tool



The "No Code" and "Other" codes were applied to segments of film where the caregiver and child were not fully engaged with each other, either due to distractions or engagements in parallel interests ("Other") or when film quality impacted coders' ability to accurately code interactions ("No Code"). "Leading" was coded during moments where the caregiver was changing the direction of play, teaching, or directing the child's attention. In contrast, "Following" behaviors occurred when caregivers allowed the child to lead the play, sitting back and supplementing the child's interests without intrusion or directions. An accompanying glossary (Figure 6) provides more expansive definitions of each behavior and outlines specific examples of dyadic behaviors aligned with each code.

Figure 6

An Expanded Glossary of Terms Relevant to the FLO Coding Tool

FLO Coding Glossary

Overall, this coding scheme aims to quantify *caregiver behaviors* within the context of dyadic free play interactions. During play, we consider the child's "spotlight" of attention. The determination of "following" and "leading" refers to whether the caregiver is in control of the child's spotlight.

This glossary is designed to complement the FLO Coding flowchart. Coders should follow coding decisions outlined by the flowchart. Definitions of each behavior are outlined in more detail below.

Following (F):

- Caregiver is attentively engaged with the child and paying attention to what the child is doing (i.e., the child is in control of the "spotlight")
    - Caregiver is sharing the child's focus.
    - Caregiver is not physically controlling the play (e.g., they may be sitting back to let the child lead)
- Caregiver is supplementing the child's activity without changing the direction of play.
    - Examples:
        - Caregiver naming something the child is doing/playing with
        - Responding to a child's serve without changing the direction of play or changing the focus of a conversation.
        - Caregiver is curious about something the child is doing (e.g., "wondering")

Leading (L):

- Caregiver is actively engaged in directing the trajectory of the child's play (i.e., the caregiver is in control of the "spotlight")
    - Caregiver may be physically controlling/manipulating the toy or the play.
    - Caregiver may be "teaching" or guiding the child's actions.
    - Caregiver may be directing the child through commands.
        - E.g., limit setting ("No") would fall into this category.
        - E.g., a child responding directly to a question or following an instruction from the caregiver is still considered leading.
    - Caregiver may change the direction of play from one thing to another.
    - Caregiver may add things to the play that the child is currently not focused on.
    - Caregiver may be diverting the child's attention away from their current interest.

Other (O):

- Caregiver is disengaged or not interacting with child; caregiver is not sharing the child's focus.
    - Examples:
        - Caregiver using their phone or distracted by something else.
        - Caregiver and child are interacting with different toys (parallel play)

Not Codable (N):

- If any part of the interaction is not visible/audible and the quality of the film makes it too difficult to make a coding decision, the coder should choose "No Code."
    - Both caregiver and child need to be in the frame
        - If either child or caregiver is off-screen it should be coded "No Code"
        - You need to be able to determine where the caregiver's attention is focused (i.e., sharing the child's focus or not) in order for a video to be considered "codeable."
            - Note: it is usually necessary for the caregiver's face/eyes to be visible in order to code, but there are some exceptions (e.g., where you can see the back of the caregiver's head but can tell what they are doing via audio or other visual cues)
    - Interaction between caregiver and child needs to be audible.
- Generally speaking, interactions that have more than two participants (i.e. multiple children or multiple caregivers) cannot be coded using the FLO Coding Protocol
    - While coding: If there is a third person present who is interacting in such a way that the caregiver's attention is shifted/changed in some way (for more than 2 seconds), then the coder should select "No Code."

At the completion of the coding process, segments marked as "No Code" were removed and percentages of Following, Leading, or Other behaviors were calculated based on the percentage of time caregivers spent engaged in each behavior within the codeable segment of video. The inter-rater reliability of double-coded films (n=53) was calculated using Noldus Observer XT to compare both the duration and sequence of coded behaviors (% agreement = 90.08%; Kappa = 0.84; Rho = 0.97).

Prior Validation of FLO.

Following the development of the FLO Coding Tool, preliminary validation analyses were conducted to evaluate the relationship between FLO behaviors and other child and caregiver outcomes. Relationships between key caregiver outcome measures and Following, Leading, and Other behaviors are outlined in detail in Table 3. There are strong associations between parents' self-reported confidence in their ability to set limits and the percentage of time they spend leading during dyadic play. In contrast, parents who endorsed confidence in their teaching abilities tended to follow their children's lead during play more often. There were no significant correlations between parent's overall reported levels of stress, exposure to adversity, or sense of competence and FLO behaviors. Notably, there were significant correlations between caregivers' ethnicity and FLO behaviors, where Hispanic caregivers were more likely to direct play and less likely to follow their children's lead than non-Hispanic caregivers.

Table 3

Correlations between Caregiver Measures and FLO

Caregiver Measures	Subscales	% Following	% Leading	% Other
Stress	PSI Parental Distress	-0.099	0.108	0.116
	PSI Dysfunctional Interaction	-0.078	0.002	0.011
	PSI Difficult Child	-0.026	-0.024	0.000
	PSI Total Stress	0.032	-0.066	0.258
	Severity of Economic Stress	-0.059	0.004	0.036
Self-Efficacy	SEPTI Nurturance	0.170	-0.065	0.278
	SEPTI Discipline/Limit Setting	0.149	0.550**	-0.329
	SEPTI Teaching	0.482*	-0.370	0.265
	SEPTI Structure/Routines	-0.061	0.082	-0.072
	SEPTI Total Score	0.105	-0.145	-0.245
Sense of Competence	PSOC	-0.055	0.089	0.205
Parent Adverse Experiences	Parent ACES	-0.021	-0.027	-0.061
Caregiver Ethnicity	Hispanic	-0.298**	0.302**	-0.115

* $p < 0.05$, ** $p < 0.001$

FLO behaviors were also compared to several child outcome measures, highlighting significant associations between Follow/Lead behaviors and child internalizing, externalizing, and language measures. Specifically, Following is associated with lower rates of child internalizing behaviors, suggesting a relationship between caregiver FLO behaviors and their children's mental health. Caregiver Leading is significantly correlated with internalizing behaviors and negatively correlated with externalizing child behaviors. Caregiver Following behaviors are also associated with higher child expressive language scores and higher levels of

child-initiated social interaction; Leading is negatively correlated with child-initiated social interaction and responses to caregiver bids for joint attention. Ultimately, these preliminary associations suggest that caregivers' Following behaviors may be related to higher child language outcomes and reduced internalizing behaviors.

Table 4

Correlations between Child Measures and FLO

Child Measures	Subscales	% Following	% Leading	% Other
Internalizing Behaviors	BITSEA Internalizing	-0.019	-0.089	0.366
	CBCL Internalizing	-0.460*	0.478**	-0.359
Externalizing Behaviors	BITSEA Externalizing	0.294	-0.356*	0.089
	CBCL Externalizing	-0.203	0.284	-0.318
	PDR Total Problems	-0.058	0.064	0.141
Child Language	PLS Auditory Comprehension	0.094	-0.062	-0.207
	PLS Expressive Communication	0.255*	-0.187	-0.199
	ESCS Initiating Behavior Request	0.085	-0.182	0.046
	ESCS Initiating Joint Attention	0.077	-0.153	-0.038
	ESCS Initiating Social Interaction	0.231*	-0.288*	0.052
	ESCS Responding to Behavior Request	0.062	-0.116	-0.032
	ESCS Responding to Joint Attention	0.177	-0.216*	0.006
	ESCS Responding to Social Interaction	0.067	-0.146	-0.051

* $p < 0.05$, ** $p < 0.001$

Data Analysis Plan and Hypotheses

Before conducting primary analyses, clinical and demographic characteristics of the sample were summarized, along with key descriptive statistics for all key study measures. A

summary of relevant means for the FIND and Control groups at baseline and post-intervention visits is provided in Results, including caregiver measures (i.e., parent sense of competence, parent self-efficacy, parent stress, parental depression and anxiety, and parent executive function), child measures (i.e., externalizing and internalizing behaviors, expressive communication, auditory comprehension, total language, and relevant domains of social communication measured by the ESCS), and dyadic measures (i.e., Child Utterances, Parent Utterances, Mean Length of Child and Parent Utterances, Conversational Turns, Following, and Leading behaviors). Repeated-measures ANOVAs were conducted to evaluate significant changes in key variables across the intervention period and significant interactions between group and time. Zero-order correlations between demographic variables and key measures are also presented in Results; these correlations were used to establish which covariates to include in relevant models. Multivariate linear regressions were used to evaluate Aim 1 and Aim 2 hypotheses; the PROCESS mediation model (Hayes, 2012) was used to evaluate Aim 3 hypotheses.

Analytic Plan for Aim 1

For Hypothesis 1, ordinary least squares regression models were run to evaluate the effects of the FIND intervention on several dyadic interaction measures: a) Child Utterances; b) Parent Utterances; c) Conversational Turns; d) parental Following behaviors; and e) parental Leading behaviors. All models run included post-intervention outcome measures as the dependent variable and baseline values, condition, and statistically indicated covariates as predictors.

Analytic Plan for Aim 2

Hypotheses relevant to Aim 2 were evaluated using ordinary least squares regression models to evaluate how post-intervention dyadic measures (e.g., post-intervention Child Utterances, post-intervention Following behaviors) predict post-intervention caregiver and child outcome measures. All models run included post-intervention outcome measures as the dependent variable and post-intervention dyadic measures, condition, and statistically indicated covariates as predictors.

Analytic Plan for Aim 3

Aim 3 hypotheses probe the possibility of mediation effects, specifically looking at how dyadic interaction variables mediate child and caregiver outcomes. Mediation effects were tested using the PROCESS macro in SPSS (Hayes, 2012), which reports regression statistics for Paths A and B, along with estimates for the direct and indirect effects. Mediation models were only tested for models that showed significant Path A and Path B results after Study Aims 1 and 2 were completed; pathways included in mediation analyses also needed to show evidence of significant intervention-related change across outcome variables, assessed during the evaluation of preliminary descriptive statistics across measures (see Tables 7, 8, 9). Mediation effects for any models with marginally significant pathways were tested post-hoc as exploratory analyses.

Covariates

Bivariate correlation analyses at baseline determined relationships between key demographic variables and measures of interest (i.e., key caregiver, child, and dyadic outcomes). Relevant covariates were included in models where they were statistically indicated; a full summary of statistically significant covariates and their relationships with key variables is included in the Results (see Tables 10, 11, 12). Potential covariates that were not significantly

correlated with key measures within a given statistical model were omitted to preserve statistical power.

Addressing Multicollinearity.

It should be noted that several analyses in the current study included both administration language and caregiver ethnicity as covariates (Hypotheses 1d-e, 2d-i, 3d-i) since these two variables were significantly associated with caregiver Following behaviors coded using the FLO Coding Tool. Although the correlation between these two measures was high (0.727, $p<0.001$), these two variables do capture important and separate information about participating families. Decisions about the inclusion and exclusion of covariates were made using a multicollinearity threshold of 0.2 and 5 for Tolerance and VIF statistics respectively (Kim, 2019). All covariates included in the models for the current study met this threshold (>0.4 Tolerance and <3 for VIF), with no covariates excluded for issues with multicollinearity.

Missing Data

The current study collected a large number of measures from caregivers and children at each study visit. Although all 92 included participants did complete both pre- and post-intervention visits, there was still a sizeable amount of missing data from participants who did not complete all questionnaires or could not complete all interactive or behavioral tasks. Table 5 outlines the final *n* for each collected measure, including the total number of missing cases per variable. Data was missing completely at random based on Little's test (Little, 1988; $\chi^2(5556) = 482.875$ p =.999). Accordingly, the decision was made not to impute missing data using single or multiple imputation techniques; all analyses were conducted using complete case analyses, omitting participants with missing data relevant to the probed hypothesis (Dettori et al., 2018; T. Little et al., 2014).

Table 5

Missing Data at Pre- and Post-Intervention Time Points

Variable	Pre-Intervention *n* (Missing)	Post-Intervention *n* (Missing)
Parent Age	92	
Child Age	92	
Child Sex	92	
Condition	92	
Language Spoken	92	
Caregiver Race	91 (1)	
Caregiver Ethnicity	92	
Parent Education	91 (1)	
Caregiver employment	92	
Household Income	91 (1)	
Parent ACES	92	
Child Adversity	88 (4)	
PSOC	92	85 (5)
SEPTI	82 (10)	81 (11)
PSI	92	87 (5)
PDR	69 (23)	69 (23)
CES-D Depression	92	87 (5)
GAD-7 Anxiety	92	87 (5)
N-Back	71 (21)	74 (18)
WCST	91 (1)	88 (4)
Internalizing Composite (BITSEA/CBCL)	64 (28)	79 (13)
Externalizing Composite (BITSEA/CBCL)	64 (28)	79 (13)
PLS	63 (29)	74 (18)
ESCS	67 (25)	62 (29)
CT (Videos Coded)	89 (3)	87 (5)
FLO (Videos Coded)	82 (10)	80 (12)

CHAPTER 3: RESULTS

Sociodemographic Sample Characteristics & Key Descriptive Statistics

Clinical and demographic characteristics of the sample are briefly summarized in this section and described in Table 6 below. Independent samples t-tests were run to compare differences in key demographic variables between the FIND and Control groups. There were significantly more boys than girls in the FIND group (M = 51.85%, SD = 0.50) compared to the Control group (M = 28.95%, SD = 0.46; t(90) = -2.224, p = 0.029). Additionally, children in the control group had experienced a higher number of adverse life events (M = 1.35, SD = 1.58) compared to children in the control group (M = 0.58, SD = 0.94; t(86) = -2.591, p = 0.011). No other demographic variables were statistically different between the two groups (p > 0.1).

Table 6

Clinical and Demographic Characteristics of the Sample

Variable	Caregiver	Child
Age (*M,* SD)	33.18 years (6.16)	24.83 months (9.69)
Gender (*n*, %)		
Female	91 (98.9%)	39 (42.4%)
Male	1 (1.1%)	53 (57.6%)
Race (*n*, %)		
Caucasian/White	76 (82.6%)	77 (83.7%)
Black or African American	6 (6.5%)	3 (3.3%)
Native American	1 (1.1%)	1 (1.1%)
Other Pacific Islander	1 (1.1%)	
Biracial/Multiracial	7 (7.6%)	10 (10.9%)
Ethnicity (*n*, %)		
Non-Hispanic White	18 (19.6%)	19 (20.7%)
Hispanic or Latinx	74 (80.4%)	72 (78.3%)

Table 6 (continued)

Variable	Caregiver	Child
Immigration Status (*n*, %)		
US-born	29 (31.5%)	
Hispanic immigrant	63 (68.5%)	
Language Spoken at Home (*n*, %)		
Only Spanish	56 (60.9%)	
Mix of Spanish and English	14 (15.2%)	
Only English	15 (16.3%)	
Parent Education		
Less than a high school degree	26 (28.6%)	
High school graduate or GED	27 (29.7%)	
Some college or technical school	14 (26.4%)	
4 year college graduate	8 (8.8%	
Master's or Doctorate degree	6 (6.6%)	
Parent Marital Status		
Never married	62 (67.4%)	
Married	11 (12.0%)	
Widowed	2 (2.2%)	
Separated	3 (3.3%)	
Divorced	14 (15.2%)	
Household Yearly Income (*M*, SD)	$27,173.68 (18639.7)	
# of people in household (*M,* SD)	2.3 (0.94)	2.75 (1.3)
% of Poverty Threshold		
50%	15 (16.3%)	
100%	42 (45.7%)	
150%	19 (20.7%)	
200%	7 (7.6%)	
>200%	7 (7.6%)	
Adverse Childhood Experiences (mean; SD)	2.22 (2.0)	1.03 (1.40)

Age, Sex, and Race/Ethnicity

The mean age for parents was 33.2 years (SD = 6.16, range 20 - 50); all primary caregivers enrolled in this study were mothers (98.9%) except for one participating father. The mean age of child participants was 30 months (SD = 9.95, range 10 - 51), with 53 boys (57.6%) and 39 girls (42.4%). The majority of caregivers identified as Caucasian/White (82.6%), with an additional 7.6% identifying as biracial or multiracial, 6.5% identifying as Black/African American, 1.1% Native American, and 1.1% Pacific Islander. Within this sample, 83.7% of children were Caucasian/White, 10.9% were biracial or multiracial, 3.3% of children were Black/African American, and 1.1% were Native American. 80.4% of caregivers and 78.3% of children in this sample identified as Hispanic or Latinx.

Immigration Status and Home Language

68.5% of caregivers in this sample identified as immigrants from other countries while only 31.5% reported being US-born. The majority of these parents (n = 58; 63.1%) immigrated from Mexico, with other parents reporting Benin, Bulgaria, Guatemala, and Peru as countries of origin. 56 (60.9%) of participating families reported using only Spanish at home, 4.3% reported using mostly Spanish, 10.9% reported using both Spanish and English, and 16.3% reported using only English. Caregivers were asked to choose their preferred language of administration for FIND intervention visits, communication with researchers, and relevant research measures; 63 (68.5%) chose to communicate in Spanish and 29 (31.5%) chose English. 34.1% of caregivers reported having received English as a Second Language (ESL) services.

Household Income & Parent Education

The mean household income was $27,173.68; notably, there was a wide range of incomes reported ($3000 - $135,000) with a significant outlier at the high end of the range ($135,000).

The average number of people supported by this income was 5.0 people (average 2.3 adults and 2.75 children per home, range 2-9 household members). 57 families (63.3%) met criteria for living under the national poverty line and 92.2% met criteria for living within 200% of the poverty line. 69.7% of caregivers reported current unemployment, with only 11.0% reporting working full time. 76.9% of primary caregivers in this study reported staying home with their children full-time.

There was a wide range of educational levels attained by parents within this sample. 13.2% did not attend high school and 15.4% reported attending some high school; 29.7% of parents graduated from high school or received a GED, 26.4% attended some college or technical school, 8.8% graduated from a 4-year college, and 6.6% had done some graduate school or received a graduate degree.

Social Welfare Benefits & Demographic Risk Factors

The majority of participating caregivers reported that they had never been married (67.4%); 51.2% of caregivers were divorced, 12.0% were married, 3.3% were separated, and 2.2% of caregivers were widowed. The majority of caregivers also endorsed other demographic risk factors: 88.0% reported currently serving as a single parent, 69.7% reported being unemployed, 38.0% reported having their first child during adolescence, 28.6% had earned less than a GED, and 18.7% reported current or past receipt of Temporary Assistance for Needy Families (TANF). Of those five identified risk factors (single parenthood, unemployment, teen parenthood, less than a GED, and TANF), 97.8% of participants endorsed at least one risk factor and 76.1% endorsed two or more risk factors.

Families participating in this study reported high use of social welfare support systems. 54.4% of families reported previous or current use of food stamps, 84.6% reported previous or

current use of WIC (Special Supplemental Nutrition Program for Women, Infants, and Children) and 12.1% were currently in welfare-supported housing. 86.8% received government-sponsored medical insurance and 26.5% reported past or current use of mental health services (e.g., psychologist, therapist, support group).

Caregiver Adversity Exposure

Parents reported on their own experiences with adverse events during their childhood using the Adverse Child Experiences questionnaire. The average number of ACES reported was 2.22; the majority of caregivers endorsed experiencing 0-3 ACES (71.7%), with 17.4% of caregivers endorsing 4-5 adverse childhood experiences and 10.9% endorsing 6 or more. The most common ACE exposures were parent separation or divorce (n = 32, 34.8%) and exposure to a substance-abusing household member (n = 31, 33.7%). A significant proportion of caregivers also reported exposure to physical abuse (n = 25, 27.2%) or sexual abuse (n = 25, 27.2%), exposure to domestic violence (n = 19, 20.7%), exposure to a household member with mental illness (n = 19, 20.7%), emotional neglect (n=19, 20.7%), physical neglect (n = 15, 16.3%), emotional abuse (n = 11, 12.0%), and having an incarcerated caregiver (n = 8, 8.8%) during their own childhood.

Child Adversity Exposure

Caregivers also completed a portion of the Life Events Checklist (LEC) on behalf of their children, which catalogued various child exposures to adverse experiences. The average number of negative life events experienced by children in this study was 1.03 (SD = 1.04), with 53.4% (n = 47) of children having experienced at least one adverse event. 28.3% of the sample (n = 26) had experienced one negative life event, 8.7% (n = 8) had experienced two negative life events, and 14.8% (n = 13) had experienced 3 or more negative life events. The majority of these

children had experienced significant family conflict (n = 36, 40.9%), including separation or divorce, exposure to family arguments, or disruption in caregiver relationships. 9.1% of the sample (n = 8) had exposure to economic stressors (e.g., food insecurity, access to basic necessities), 9.1% (n = 8) of the sample had exposure to violence, and 1.1% (n = 1) had exposure to discrimination-based stressors. Given the ages of the children in this study (< 3 years), this variable is also an indication of recent family experiences of conflict, economic stressors, violence, and other stressors.

Comparison of Key Study Variables at Pre- and Post-Intervention

The means and standard deviations for key outcome measures are summarized by group (FIND vs. Control group) at pre- and post-intervention time points and presented in Tables 7, 8, and 9. Independent sample t-tests were conducted to test for significant differences between groups at the pre-intervention time point; t and p values for these between-group differences are also listed in Tables 7, 8, and 9. Repeated-measures ANOVAs were conducted to evaluate whether group means changed significantly across the intervention period; F and p values highlighting intervention effects (Time*Condition interaction) are listed in each corresponding table. Further analysis of change in outcome measures over time will be presented according to each study aim and organized by hypothesis throughout the results section.

Table 7

Comparison of Caregiver Group Means at Pre- and Post- Intervention

Variable	Pre-Intervention Mean (SD)		Baseline Group Differences		Post-Intervention Mean (SD)		Intervention Effects (ANOVA)	
	FIND	Control	t	*p*	FIND	Control	F	*p*
PSOC	39.22 (4.63)	40.39 (4.43)	1.217	0.227	41.14 (4.70)	40.44 (4.53)	4.750	0.032*
SEPTI Total	153.87 (17.12)	151.51 (19.01)	-0.589	0.558	163.45 (2.48)	158.30 (2.79)	1.233	0.271
SEPTI Nurturance	42.21 (3.46)	41.03 (3.92)	-1.45	0.152	42.09 (4.06)	42.20 (3.55)	0.733	0.395
SEPTI Discipline/ Limit Setting	28.11 (7.34)	27.77 (7.26)	-0.205	0.838	31.67 (6.86)	30.31 (6.84)	0.655	0.421
SEPTI Teaching	45.30 (5.99)	45.83 (5.12)	0.422	0.674	48.67 (5.13)	46.80 (6.07)	5.435	0.023*
SEPTI Instrumental Care/Routines	38.25 (6.98)	36.88 (8.64)	-0.794	0.430	39.61 (6.53)	39.42 (7.13)	0.299	0.586
PSI Total	70.50 (16.55)	71.84 (19.83)	0.353	0.725	64.37 (22.62)	65.75 (20.02)	0.015	0.904
PSI Parental Distress	27.48 (6.84)	27.84 (8.79)	0.221	0.825	23.77 (9.39)	24.94 (8.98)	0.140	0.709
PSI Dysfunctional Interaction	18.17 (5.15)	18.74 (6.09)	0.485	0.629	18.471 (8.28)	17.58 (6.11)	0.850	0.359
PSI Difficult Child	24.85 (8.88)	25.26 (8.05)	0.227	0.821	22.14 (8.19)	23.22 (7.89)	0.037	0.849
PDR Daily Stress	1.37 (1.20)	1.28 (0.84)	-0.313	0.755	1.17 (1.02)	1.37 (1.25)	0.013	0.911
PDR Sum of Daily Hassles	6.36 (6.66)	7.76 (6.48)	0.845	0.401	6.14 (4.83)	7.13 (5.07)	0.574	0.452
PDR Negative Mood	8.00 (7.07)	8.36 (7.54)	0.198	0.843	4.67 (5.10)	6.50 (8.51)	0.907	0.346
PDR Positive Mood	9.23 (2.92)	9.66 (2.15)	0.647	0.520	16.97 (6.89)	16.37 (6.30)	0.001	0.980
CES-D (Depression)	10.07 (6.68)	11.21 (9.73)	0.665	0.508	10.65 (7.98)	12.66 (8.69)	1.15	0.286
GAD-7 (Anxiety)	5.19 (4.01)	5.21 (4.81)	0.027	0.978	4.83 (4.18)	3.54 (3.28)	2.068	0.154
N-Back % Accuracy 2-back	83.78 (7.44)	83.93 (8.29)	-0.074	0.941	85.01 (7.59)	84.72 (9.62)	0.325	0.571

Table 7 (continued)

Variable	Pre-Intervention Mean (SD)		Baseline Group Differences		Post-Intervention Mean (SD)		Intervention Effects (ANOVA)	
	FIND	Control	t	*p*	FIND	Control	F	*p*
N-Back % Accuracy 3-back	80.21 (7.40)	79.75 (5.83)	-0.284	0.777	79.14 (8.45)	78.42 (7.21)	0.009	0.924
N-Back % Accuracy Total	86.44 (6.74)	87.08 (5.76)	0.423	0.674	87.18 (6.01)	86.51 (6.67)	1.717	0.195
N-Back Hits/False Alarms	6.88 (8.15)	5.69 (4.49)	-0.713	0.239	6.46 (5.43)	8.02 (9.72)	0.066	0.798
WCST Trials Completed	104.78 (29.91)	102.73 (32.21)	-0.311	0.757	98.60 (33.38)	96.00 (35.62)	0.024	0.876
WCST % Accuracy	41.22 (12.03)	39.85 (14.57)	-0.492	0.312	46.52 (15.48)	46.20 (14.42)	0.174	0.678
WCST % Perseverative	8.41 (7.09)	8.38% (6.74)	-0.019	0.492	7.74 (5.44)	8.26 (8.43)	0.114	0.736
WCST % Conceptual	24.49 (17.01)	24.29% (20.20)	-0.051	0.959	29.32 (20.82)	30.97 (22.00)	0.153	0.697

Caregiver Self-Efficacy

PSOC.

At the pre-intervention visit, the mean caregiver PSOC score was 39.22 (SD = 4.63) for caregivers in the FIND group and 40.39 (SD = 4.43) for caregivers in the control group. At baseline, these between-group means were not significantly different from each other ($t(90) = 1.217$, $p = 0.227$). At post-intervention, FIND caregivers had a mean PSOC score of 41.14 (SD = 4.70) and caregivers in the control group had a mean score of 40.44 (SD = 4.53). This change reflects a statistically significant increase for the FIND group compared to the control group over time ($F(85) = 4.750$, $p = 0.032$).

SEPTI-TS.

Caregiver self-efficacy was measured using four subscales of the SEPTI-TS (Nurturance, Discipline/Limit Setting, Teaching, and Instrumental Care/Routines) in addition to the SEPTI-TS total score. At pre-intervention, the total SEPTI score was 153.87 (SD = 17.12) for the FIND group and 151.51 (SD = 19.01) for the control group. These pre-intervention scores – along with all SEPTI subscale scores at the pre-intervention time point – were not significantly different between the FIND and Control group (see Table 7). At the post-intervention visit, SEPTI total scores had increased for both groups: FIND caregivers had a mean score of 163.45 (SD = 2.48) and caregivers in the control group had a mean score of 158.30 (SD = 2.79). While the overall increase in SEPTI total scores over time was significant ($F(73) = 14.75$, $p < 0.001$), this change was not significantly different between the FIND and control groups ($F(73) = 0.437$, $p = 0.511$). A similar pattern occurred with the SEPTI Discipline/Limit Setting subscale, where pre-intervention scores (FIND: $M = 28.11$, SD =7.34; Control: $M = 27.77$, SD = 7.26) showed a significant increase over time (Post FIND: $M = 31.67$, SD = 6.86; Post Control: $M = 30.31$, SD = 6.84; $F(73) = 12.91$, $p < 0.001$), but no significant interaction between time and condition ($F(73) = 0.655$, $p = 0.421$). For the SEPTI Teaching subscale, caregivers in the FIND group showed a significant increase over time (FIND Pre: $M = 45.30$, SD = 5.99; FIND Post: $M = 48.67$, SD=5.13) compared to the control group (Control Pre: $M = 45.83$, SD = 5.12; Control Post: $M = 46.80$, SD = 6.07); $F(73) = 5.435$, $p = 0.023$). There were no significant differences in pre-post means for the SEPTI Nurturance or SEPTI Instrumental Care/Routines subscales.

Caregiver Stress

PSI.

Caregiver stress was measured using three subscales of the PSI (Parental Distress, Parent-Child Dysfunctional Interaction, and Difficult Child) in addition to the PSI total score. At pre-intervention, the mean total score was 70.50 (SD = 16.55) for caregivers in the FIND condition and 71.84 (SD = 19.83) for caregivers in the control condition. These two means were not statistically different ($t(90) = 0.353$, $p = 0.725$), suggesting no meaningful differences between groups at baseline; further independent samples t-tests for all subscales of the PSI at baseline indicate no significant differences between groups on any of the PSI subscales. At post-intervention, both groups showed a significant decrease in PSI Total Scores (FIND Post: M = 163.45, SD = 2.48; Control Post: M = 158.30, SD=2.79; $F(85) = 7.966$, $p = 0.006$); however, the Time *x* Condition interaction was not significant ($F(85) = 0.015$, p =0.904).

PDR.

In addition to the PSI, four items from the Parent Daily Report were used to evaluate parent stress at pre- and post-intervention timepoints: Daily Stress, Total Daily Hassles, Total Negative Mood, and Total Positive Mood. For all four of these items, there were no significant differences between FIND and Control group means at baseline. Daily Stress ratings remained consistent from pre-intervention (FIND: M = 1.37, SD = 12.0; Control: M = 1.28, SD = 0.84) to post-intervention (FIND: M = 1.17, SD = 1.02); Control: M = 1.37, SD = 1.25). Total Daily Hassles also remained stable from pre-intervention (FIND: M = 6.36, SD = 6.66; Control: M = 7.76, SD = 6.48) to post-intervention (FIND: M = 6.14, SD = 4.83; Control: M = 7.13, SD = 5.07). Mean Total Negative Mood scores showed a significant decrease between pre-intervention (FIND: M = 8.00, SD = 7.07; Control: M = 8.36, SD = 7.54) and post-intervention (FIND: M =

4.67, SD = 5.10; Control: M = 6.50, SD = 8.51; F(47) = 4.459, p = 0.040), although this was a main effect for time only; the FIND group did not show a significant increase when compared to the Control group over time (F(47) = 0.907, p = 0.346). In a similar pattern, the mean Total Positive Mood scores showed a significant increase between pre-intervention (FIND: M = 9.23, SD = 2.92; Control: M = 9.66, SD = 2.15) and post-intervention (FIND: M = 16.97, SD = 6.89; Control: M = 16.37, SD = 6.30). This was a significant increase in scores over time (F(47) = 45.562, p < 0.001), but there were no significant differences in the interaction between time and group assignment (F(47) = 0.001, p = 0.980).

Caregiver Mental Health

CES-D.

At the pre-intervention visit, the mean depression score for caregivers in the FIND group was 10.07 (SD = 6.68) while caregivers in the Control group had a mean score of 11.21 (SD = 9.73); these scores were not significantly different at baseline (t(90) = 0.665, p=0.508). Post-intervention, CES-D scores remained stable for caregivers in the FIND group (M = 10.65, SD = 7.98) and rose slightly for those in the Control group (M = 12.66, SD = 8.69), although this increase was not statistically significant (F(85) = 1.15, p = 0.286).

GAD-7.

At the pre-intervention visit, the mean anxiety scores for caregivers in the FIND group (M = 5.19, SD = 4.01) and Control group (M = 5.21, SD = 4.81) were not significantly different (t(90) = 0.027; p = 0.978). These scores remained fairly consistent at post-intervention, with a slightly larger decrease in anxiety for caregivers in the Control group (M = 3.54, SD = 3.28) compared to FIND (M = 4.83, SD = 4.18). This slight decrease in anxiety scores over time is

marginally significant (F(85) = 3.216, p =0.077); the interaction between time and group assignment is not significant (F(85) =2.068, p = 0.154).

Caregiver Executive Function

N-Back.

Four variables from the N-back task were used to evaluate caregiver working memory, attention, and cognitive control: Percent Accuracy at 2-back, Percent Accuracy at 3-back, Overall Percent Accuracy, and the Ratio of Hits to False Alarms. There were no significant differences between FIND and Control group means across all four of these measures at baseline. Caregiver accuracy remained fairly consistent from pre-intervention to post-intervention across the 2-back (pre- FIND: M = 83.78%, SD = 7.44; post-FIND: M = 85.01%, SD = 7.59; pre-Control: M = 83.93%, SD = 8.29; post-Control: M = 84.72%, SD = 9.72), 3-back (pre- FIND: M = 80.21%, SD = 7.40; post-FIND: M = 79.14%, SD = 8.45; pre-Control: M = 79.75%, SD = 5.83; post-Control: M = 78.42%, SD = 7.21), and overall (pre- FIND: M = 86.44%, SD = 6.74; post-FIND: M = 87.18%, SD = 6.01; pre-Control: M = 87.08%, SD = 5.76; post-Control: M = 86.51%, SD = 6.67). The ratio of hits to false alarms remained stable for the FIND group (pre- FIND: M = 6.88, SD = 8.15; post-FIND: M = 6.46, SD = 5.43) and increased slightly for the Control group (pre- Control: M = 5.69, SD = 4.49; post- Control: M = 8.02, SD = 9.72). There were no significant time or Condition x Time effects across all measured N-back variables (p > 0.1).

Wisconsin Card Sort Task.

This study used four measured variables from the Wisconsin Card Sort Task to evaluate cognitive flexibility, attentional control, problem-solving, and other executive functions: Total Number of Trials Completed, % Accuracy across Trials, Perseverative Response Rate, and

Conceptual Response Rate. There were no significant differences between FIND and Control group scores at baseline across any of these four measured domains ($p > 0.3$). Total trials completed decreased across the intervention period for both the FIND (pre: $M = 104.78$, SD = 29.91; post: $M = 98.60$, SD = 33.38) and Control groups (pre: $M = 102.73$, SD = 32.21; post: $M = 96.00$, SD = 35.62); this decrease in scores over time was marginally significant ($F(85) = 2.970$, $p = 0.088$), although there was not a significant difference between declining scores in the FIND vs. the Control group over time ($F(85) = 0.024$, $p = 0.876$). Overall percent accuracy on the WCST increased significantly across the intervention period for both the FIND (pre: $M = 41.22\%$, SD = 12.03; post: $M = 46.52\%$, SD = 15.48) and Control groups (pre: $M = 39.85\%$, SD = 14.57; post: $M = 46.20\%$, SD = 14.42; $F(85) = 17.159$, $p < 0.001$). This difference was not significant when considering the impact of FIND compared to the Control group over time ($F(85) = 0.174$, $p = 0.678$). The perseverative response rate remained consistent for both groups across the intervention period (pre- FIND: $M = 8.41\%$, SD = 7.09; post- FIND: $M = 7.74\%$, SD = 5.44; pre- Control: $M = 8.38\%$, SD = 6.74; post- Control: $M = 8.26\%$, SD = 8.43). Finally, the percentage of Conceptual Responses increased significantly for both the FIND (pre: $M = 24.49\%$, SD = 17.01; post: $M = 29.32\%$, SD = 20.82) and Control groups over time (pre: $M = 24.29\%$, SD = 20.2; post: $M = 30.97\%$, SD = 22.00; $F(85) = 1.965$, $p = 0.006$); as with the other measured domains of the WCST, this effect was significant for time but did not reflect significant group differences between the FIND and Control group across the intervention period ($F(85) = 0.153$, $p = 0.697$).

Child Behaviors and Socio-Emotional Outcomes

PDR.

Child problematic behaviors were measured using the PDR Total Behaviors score. At baseline, mean scores for the FIND (M = 12.64, SD = 9.17) and Control (M = 14.98, SD =10.66) groups were not significantly different (t(58) = 0.904, p = 0.370). At the post-intervention timepoint, the FIND group scores stayed consistent (M = 12.05, SD = 11.12), while Control group scores decreased (M = 9.84, SD = 8.32). These decreases reflect a significant change in scores across time (F(44) = 6.37, p = 0.015), although the interaction between time and group condition was not significant (F(44) = 0.052, p = 0.821).

CBCL/BITSEA Composite.

The BITSEA/CBCL composite scores for Internalizing and Externalizing reflect a percentage of endorsed behaviors from the total possible number of questions on each scale (BITSEA and CBCL). This percentage was calculated across participants who completed the BITSEA or CBCL measures, since these two questionnaires were given differentially based on child age. At the pre-intervention timepoint, caregivers in the FIND condition endorsed 14.87% (SD = 14.37) of possible internalizing behaviors while caregivers in the Control condition endorsed 12.19% (SD = 8.64). This baseline difference was not statistically significant (t(62) = -0.827, p = 0.206). Post-intervention, both FIND and Control groups endorsed fewer internalizing behaviors (FIND: M = 9.99, SD = 8.04; Control: M = 9.01, SD = 7.30). This is a statistically significant decrease over time (F(50) = 9.59, p = 0.003), although there was not a significant difference between the FIND and Control group over time (F(50) = 0.152, p = 0.699).

For the BITSEA/CBCL Externalizing composite scores, caregivers in the FIND group endorsed slightly more externalizing behaviors (M = 28.29%, SD = 18.55) than caregivers in the

Control group (M = 24.83%, SD = 17.74) at baseline; this difference in baseline scores is not statistically significant (t(62) = -0.734, p = 0.465). At the post-intervention time point, caregivers in the FIND condition endorsed a lower percentage of possible externalizing behaviors (M = 23.81%, SD = 16.29) while caregivers in the Control group endorsed a higher percentage of possible externalizing behaviors (M = 27.80%, SD = 18.51); these differences in the FIND and Control group means over time are not statistically significant (F(50) = 0.000, p = 0.995).

Child Language

PLS-5.

Three scores from the Preschool Language Scales (PLS-5) were used to evaluate children's language proficiency: PLS Expressive Communication, Auditory Comprehension, and Total Language, which represents the sum total of the EC and AC subscales. At baseline, there was a significant difference between the FIND (M = 90.09, SD =22.24) and Control groups (M = 100.53, SD = 13.34; t(61) = 2.232, p = 0.029). Across the intervention period, children's scores in the FIND condition increased (M = 101.98, SD = 18.06), while children's scores in the Control condition remained consistent with their pre-intervention performance (M = 99.67, SD = 13.21). This difference in scores between the FIND and Control groups over time is marginally significant (F(54)=3.46, p = 0.068). For the Auditory Comprehension subscale, baseline scores for the FIND (M = 100.15, SD = 15.52) and Control groups (M = 103.20, SD = 11.54) did not differ significantly (t(62) = 0.883, p = 0.381). Across the intervention period, scores from children in the FIND group increased slightly (Post M = 103.89, SD = 16.71) while scores from children in the Control group decreased (Post M = 98.51, SD = 12.14). These differences between the FIND and Control groups over time were marginally significant (F(57) = 3.66, p = 0.61). Scores for the PLS Total Language score followed a similar pattern; at baseline, the FIND

group (M = 95.84, SD = 13.67) had significantly lower scores than the Control group (M = 102.07, SD = 11.53; t(60) = 1.94, p = 0.029). Across the intervention period, children's scores from the FIND group increased (Post M = 103.36, SD =18.27) while scores from the Control group decreased slightly (Post M = 99.15, SD = 11.27). This interaction between time and assigned group condition was significant (F(53) = 4.011, p = 0.050).

ESCS.

For the Early Social Communication Scales, three primary interaction behaviors were scored: joint attention, behavior request, and social interaction. For each behavior, children were evaluated both in terms of how often they *initiated* the behavior and how often they *responded* to caregivers' bids in that domain. Across five of the six subscales, FIND and Control group means were not significantly different from each other at baseline (see Table 8 for a full summary of statistical tests). For the Responding to Social Interaction subscale, the FIND group (M = 5.33, SD = 2.15) had significantly lower baseline scores than the Control group (M = 6.54, SD = 2.34; t(65) = 2.15, p = 0.018). Across the intervention period, children in the FIND group's scores on this subscale increased (Post M = 6.95, SD = 2.63) while scores in the Control group decreased slightly (Post M = 6.04, SD = 2.23). This difference between the FIND and Control group's scores over time was statistically significant (F(61) = 8.758, p = 0.004). A similar pattern emerged for the other Social Interaction subscale (*Initiating* Social Interaction); there was a marginally significant difference between the FIND and Control group's scores over time (F(61) = 3.211, p = 0.078), with FIND scores showing a slight increase over the intervention period (Pre M = 2.00, SD = 1.09; Post M = 2.30, SD = 1.03) and Control group scores decreasing across the same time period (Pre M = 2.30, SD = 1.03; Post M = 2.00, SD = 1.13).

Significant changes across the intervention period were also present for the Joint Attention subscale. Scores increased across both FIND and Control groups for the Responding to Joint Attention means (Pre FIND: M = 6.49, SD = 3.13; Post FIND: M = 8.07, SD = 3.52; Pre Control: M = 5.92, SD = 3.83; Post Control: M = 8.11, SD = 3.52). These increases show a significant effect of time ($F(61) = 13.61$, $p < 0.001$), although there were no significant differences between the FIND and Control groups over time ($F(61) = 0.045$, $p = 0.833$). For the Initiating Joint Attention subscale, FIND scores increased slightly over the intervention period (Pre M = 13.37, SD = 6.96; Post M = 14.18, SD = 7.67) while Control scores decreased (Pre M = 14.29, SD = 7.69; Post M = 9.91, SD = 6.39). These changes reflect a significant interaction between group and time ($F(61) = 7.553$, $p = 0.008$). Scores on the Behavior Request subscale stayed consistent across the intervention period and were not meaningfully different between groups (for all reported means and statistical tests, see Table 8).

Table 8

Comparison of Group Means at Pre- and Post-Intervention for Key Child Outcomes

Variable	Pre-Intervention Mean (SD)		Baseline Group Differences		Post-Intervention Mean (SD)		Intervention Effects (ANOVA)	
	FIND	Control	t	*p*	FIND	Control	F	*p*
PDR Total	12.64 (9.17)	14.98 (10.66)	0.904	0.370	12.05 (11.12)	9.84 (8.32)	0.052	0.821
BITSEA/CBCL Composite: Internalizing	14.87 (14.37)	12.19 (8.64)	-0.827	0.206	9.99 (8.04)	9.01 (7.30)	0.152	0.699
BITSEA/CBCL Composite: Externalizing	28.29 (18.55)	24.83 (17.74)	-0.734	0.465	23.81 (16.29)	27.80 (18.51)	0.000	0.995
PLS Expressive Communication	90.09 (22.24)	100.53 (13.34)	2.232	0.029*	101.98 (18.06)	99.67 (13.21)	3.46	0.068
PLS Auditory Comprehension	100.15 (15.52)	103.20 (11.54)	0.883	0.381	103.89 (16.71)	98.51 (12.14)	3.66	0.061
PLS Total Language	95.84 (13.67)	102.07 (11.53)	1.94	0.029*	103.36 (18.27)	99.15 (11.27)	4.011	0.050*

Table 8 (continued)

Variable	Pre-Intervention Mean (SD)		Baseline Group Differences		Post-Intervention Mean (SD)		Intervention Effects (ANOVA)	
	FIND	Control	t	*p*	FIND	Control	F	*p*
ESCS Initiating Joint Attention	13.37 (6.96)	14.29 (7.69)	0.499	0.619	14.18 (7.67)	9.91 (6.39)	7.553	0.008**
ESCS Responding to Joint Attention	6.49 (3.13)	5.92 (3.83)	-0.662	0.510	8.07 (3.52)	8.11 (3.52)	0.045	0.833
ESCS Initiating Behavior Request	13.26 (7.12)	11.92 (6.60)	-0.757	0.452	13.65 (6.83)	12.70 (7.14)	0.004	0.952
ESCS Responding to Behavior Request	9.07 (5.72)	8.29 (4.57)	-0.572	0.569	9.28 (4.92)	9.13 (4.47)	0.000	0.999
ESCS Initiating Social Interaction	2.00 (1.09)	2.21 (1.23)	0.703	0.485	2.30 (1.03)	2.00 (1.13)	3.211	0.078
ESCS Responding to Social Interaction	5.33 (2.15)	6.54 (2.34)	2.15	0.018*	6.95 (2.63)	6.04 (2.23)	8.758	0.004**

Dyadic Interaction Measures

CT.

Five relevant subscales from the CT measure are reported in Table 9: number of Child Utterances, mean length of Child Utterances, number of Parent Utterances, mean length of Parent Utterances, and the number of Conversational Turns. There were no significant baseline differences between groups across any of these five metrics (see Table 9 for significance tests). At the pre-intervention timepoint, the mean number of Child Utterances during dyadic play was 38.81 (SD = 28.90) for children in the FIND group and 45.46 (SD = 29.27) for children in the Control group. Post-intervention means were higher for both groups (FIND: M = 47.77, SD = 29.49; Control: M = 50.49, SD = 36.31) and there was a statistically significant increase in Child Utterances over time ($F(84) = 4.415$, $p = 0.039$). This increase in Child Utterances did not differ significantly between the FIND and Control groups ($F(84) = 0.157$, $p = 0.693$). The mean length

of Child Utterances was not significantly different between groups or time points (F(84) = 0.069, p = 0.794). Across the intervention period, the number of Parent Utterances decreased between pre-intervention (FIND: M = 97.47, SD = 37.92; Control: M = 108.00, SD = 28.22) and post-intervention (FIND: M = 93.48, SD = 30.49; Control: M = 97.66, SD = 35.93) timepoints; this decrease over time was marginally significant (F(84) = 3.352, p = 0.071), but there was not a significant difference between FIND and Control group changes over time (F(84) = 0.434, p = 0.512). There were not notable differences between the mean length of Parent Utterances between groups or across time points (see Table 9 for all mean scores). The FIND group averaged 51.77 (SD = 42.03) Conversational Turns at the pre-intervention time point while the Control group averaged 62.47 (SD = 41.31). While the mean Conversational Turns scores increased slightly for both groups (Post FIND M = 57.15, SD = 34.21; Post Control M = 63.14, SD = 43.50), there was significant variance in Conversational Turns across the sample and these small increases were not significant for time (F(84) = 0.652, p = 0.422) or the interaction between group condition and time (F(84) = 0.078, p = 0.781).

FLO.

Three key metrics were assessed for the FLO Coding Tool: the percentage of time that caregivers engaged in Following, Leading and Other behaviors. At baseline, caregivers in the FIND group followed their child's lead during play 35.03% of the time (SD = 19.87) while caregivers in the Control group engaged in Following behaviors 45.74% of the time (SD = 21.89). This baseline difference was statistically significant (t(80) = 2.295, p = 0.024). At the post-intervention timepoint, FIND parents increased the percentage of time they followed their child's lead during play to 51.91% (SD = 23.62); Parents in the control group showed a modest increase in Following behaviors over time (Post M = 49.00, SD = 26.79). This difference reflects

a statistically significant interaction between time and group condition (F(71) = 8.148, p =0.006). The opposite pattern appeared with Leading behaviors: there was a significant decrease in caregiver Leading across the intervention period between the FIND (Pre M = 62.65, SD = 20.79; Post M = 45.48, SD = 23.68) and Control groups (Pre M = 53.59, SD = 22.46; Post M = 49.98, SD = 27.10; F(71) = 6.723, p = 0.012). As with Following, the difference between the FIND and Control group baseline Leading means was statistically significant (p(80) = -1.874, p = 0.065). There were no significant differences between the FIND and Control group Other behaviors across the intervention period (F(13) = 0.019, p=0.892), although it should be noted that Other behaviors occurred much less frequently than Leading or Following behaviors. The mean percentage of time that caregivers engaged in Other behaviors for the FIND group was 5.14% (SD = 4.95) at baseline and 5.95% (SD = 6.21) at post-intervention. For the control group, the mean percentage of time that caregivers engaged in Other behaviors was 3.15% (SD=3.45) at baseline and 4.98% (SD=3.53) at post-intervention. Notably, only 13 of the 92 caregiver/child dyads engaged in interactions that contained any Other scores during both pre- and post-visits.

Table 9

Comparison of Group Means at Pre- and Post-Intervention for Key Dyadic Outcomes

Variable	Pre-Intervention Mean (SD)		Baseline Group Differences		Post-Intervention Mean (SD)		Intervention Effects (ANOVA)	
	FIND	Control	t	*p*	FIND	Control	F	*p*
CT Total Child Utterances	38.81 (28.90)	45.46 (29.27)	1.075	0.2285	47.77 (29.49)	50.49 (36.31)	0.157	0.693
CT Mean Length of Child Utterances	777.69 (256.02)	760.01 (257.43)	-0.319	0.750	764.57 (246.04)	722.06 (316.80)	0.069	0.794
CT Total Parent Utterances	97.47 (37.92)	108.00 (28.22)	1.419	0.159	93.48 (30.49)	97.66 (35.93)	0.434	0.512
CT Mean Length of Parent Utterances	1117.02 (430.45)	1064.67 (261.78)	-0.652	0.258	1052.44 (287.04)	1092.54 (463.71)	1.214	0.274
CT Conversational Turns	51.77 (42.03)	62.47 (41.31)	1.187	0.239	57.15 (34.21)	63.14 (43.50)	0.078	0.781

Table 9 (continued)

Variable	Pre-Intervention Mean (SD)		Baseline Group Differences		Post-Intervention Mean (SD)		Intervention Effects (ANOVA)	
	FIND	Control	t	*p*	FIND	Control	F	*p*
FLO % Follow	35.03 (19.87)	45.74 (21.89)	2.295	0.024*	51.91 (23.62)	49.00 (26.79)	8.148	0.006**
FLO % Lead	62.65 (20.79)	53.59 (22.46)	-1.874	0.065	45.48 (23.68)	48.98 (27.10)	6.723	0.012*
FLO % Other	5.14 (4.95)	3.15 (3.45)	-0.987	0.332	5.95 (6.21)	4.98 (3.53)	0.019	0.892

Zero-Order Bivariate Correlation Analyses

Zero-order correlations were calculated between all study variables and socio-demographic variables that could be potential confounds, including parent and child age, child sex, group condition, home language, caregiver ethnicity, parent education, household income, parent ACEs, and child adversity exposure. Tables 10, 11, and 12 show the zero-order bivariate correlations for demographic variables and the caregiver, child, and dyadic measures used in the current study. Independent t-tests were used to determine significant associations between dichotomous variables (e.g., child sex, group condition, language spoken, ethnicity). Variables with statistically significant ($p < 0.05$) associations are summarized and discussed in text below.

Correlations between Demographic Variables and Caregiver Measures

Parent age was significantly associated with that ratio of hits to false alarms on the N-back task ($r = 0.276$, $p = 0.022$) and marginally associated with overall percent accuracy on the N-back task ($r = 0.228$, $p = 0.056$). Child age was positively associated with several variables that measured parent stress. Specifically, parents with older children tended to report higher levels of stress via the PSI Total Stress score ($r = 0.267$, $p = 0.01$), the PSI Difficult Child subscale ($r = 0.333$, $p = 0.001$), and on daily ratings of stress level taken using the PDR ($r =$

0.295, p = 0.016). Child age was also marginally associated with caregiver performance on the 3-back task (r = 0.209, p = 0.081).

Child sex was significantly associated with PDR Daily Stress levels (t(64) = -2.064, p = 0.043), with parents of girls reporting higher levels of stress (M = 1.64, SD = 1.162) than parents of boys (M = 1.11, SD = 0.953).

Caregiver-preferred language (i.e. language of administration for all intervention and research visits) was significantly associated with caregiver performance across the N-back and Wisconsin Card Sort Tasks. Specifically, caregivers who conducted visits in English scored significantly higher on 2-back trials (M = 89.73%, SD = 7.44) compared to Spanish-speakers (M = 82.98%, SD = 8.01; t(72) = 3.33, p < 0.001); English-speakers also performed better across all N-back trials (English: M = 89.99%, SD = 5.66; Spanish: M = 85.24%, SD = 6.07; t(69) = 3.11, p = 0.001) and had a higher ratio of hits to false alarms (English: M = 9.78, SD = 9.92; Spanish: M = 4.89, SD = 4.28; t(67) = 2.88, p = 0.003). Home language was marginally associated with overall accuracy on the Wisconsin Card Sort Task (English: M = 44.47%, SD = 15.37; Spanish: M = 38.97%, SD = 11.64; t(89) = 1.88, p = 0.063) and the overall % of Conceptual Responses (English: M = 29.52%, SD = 21.83; Spanish: M = 22.13%, SD = 16.16; t(89) = 1.80, p =0.075).

Caregiver ethnicity was also significantly correlated with several measures of executive function captured by the N-back and Wisconsin Cart Sort tasks. Latino/a caregivers scored significantly lower on the 2-back (Non-Latino/a: M = 91.67%, SD = 4.66; Latino/a: M = 83.57%, SD = 8.33; t(72) = 3.25, p = 0.002) and overall N-back tasks (Non-Latino/a: M = 90.37%, SD = 6.24; Latino/a: M = 85.89%, SD = 6.08; t(69) = 2.39, p = 0.019). Latino/a caregivers also had significantly lower accuracy scores on the Wisconsin Card Sort Task (Non-Latino/a: M = 47.73%, SD = 15.41; Latino/a: M = 38.92%, SD = 11.90; t(89) = 2.65, p = 0.010), including

lower Conceptual Response scores (Non-Latino/a: M = 34.46%, SD = 10.88; Latino/a: M = 21.93%, SD = 16.85; t(89) = 2.69, p = 0.008). On average, Latino/a caregivers completed more trials (M = 107/49, SD = 28.54) than their non-Latino/a counterparts (M = 89.56, SD = 35.61; t(89) = -2.27, p =0.026).

Parent education level was negatively associated with Dysfunctional Interaction (r = -0.215, p = 0.041) and positively associated with PDR ratings of Daily Stress (r = 0.258, p = 0.038). Parental education was also significantly correlated with several measures of caregiver executive function. Caregiver education was negatively correlated with accuracy on the 2-back (r = -0.361, p = 0.02) and 3-back trials (r = -0.370, p = 0.001) and the total number of Wisconsin Card Sort Trials completed (r = -0.299, p = 0.004). Education was positively correlated with overall N-back accuracy (r = 0.443, $p < 0.001$), Wisconsin Card Sort Task overall accuracy (r = 0.284, p = 0.007), and the accuracy of Conceptual Responses on the WSCT (r = 0.251, p = 0.017).

Household income was positively associated with caregivers' endorsed confidence in their ability to care for their children and complete daily routines (SEPTI Instrumental Care; r = 0.279, p = 0.012). Group condition was not significantly associated with any of the key caregiver variables ($p > 0.1$).

Beyond demographic data, caregiver ACES were significantly associated with a number of key caregiver outcome measures. Caregiver ACE scores were negatively associated with parent sense of competence (PSOC; r = -0.403, $p < 0.001$), overall parent self-efficacy (SEPTI Total; (r = -0.280, p = 0.011), self-perceived caregiver nurturance (r = -0.233, p = 0.044), and limit setting (r = -0.254, p = 0.021). Caregiver ACES were also positively correlated with depression (CES-D; r = 0.275, p = 0.008) and anxiety (GAD-7; r = 0.414, $p < 0.001$). There

was a marginally significant positive correlation between Caregiver ACES and parents' perception of daily events as stressful (PDR Daily Hassle; $r = 0.236$, $p = 0.051$).

Child exposure to significant life events also proved to be correlated with a number of key caregiver outcome measures. Since the events probed highlight recent family life stressors (e.g., family conflict, economic stressors, violence), this variable can be considered a proxy for recent stressful events impacting both caregivers and children enrolled in this study. Child adverse experiences was significantly correlated with parental distress (PSI; $r = 0.213$, $p = 0.046$), parental anxiety (GAD-7; $r = 0.223$, $p = 0.037$), and PDR ratings of daily hassles ($r = 0.256$, $p = 0.037$) and negative mood $r = 0.319$, $p = 0.008$). Adverse child life events were also significantly correlated with all measured domains of caregiver executive function via the N-back ($r = 0.253$ - 0.300, $p = 0.014$ - 0.039). and Wisconsin Card Sort Task ($r = 0.242$ - 0.334, $p = 0.001$ - 0.023).

Table 10

Correlations between Demographic Variables and Baseline Caregiver Measures

Variable	Parent Age	Child Age	Child Sex	Group Condition	Language (Admin.)	Ethnicity	Parent Education	Household Income	Parent ACES	Child Adverse Events
PSOC	-0.115	-0.170	-0.124	-0.127	0.126	0.119	-0.191	-0.015	-0.403**	-0.087
SEPTI Total	-.137	0.015	-0.013	0.066	-0.032	0.062	0.113	0.189	-0.280*	-0.020
SEPTI Nurturance	0.092	0.000	0.055	0.160	0.005	-0.019	0.037	0.188	-0.223*	0.010
SEPTI Teaching	-0.053	0.104	-0.094	-0.047	-0.057	0.043	0.045	0.168	-0.172	-0.085
SEPTI Instrum. Care	0.197	-0.001	-0.021	0.088	-0.070	0.057	0.152	0.279*	-0.177	0.040
SEPTI Limit Setting	0.122	-0.041	0.035	0.023	0.037	0.070	0.064	-0.055	-0.254*	-0.030
PSI Total Stress	0.142	0.267*	-0.013	-0.037	-0.005	0.142	-0.035	-0.088	0.185	0.080
PSI Parental Distress	0.114	0.133	-0.054	-0.023	-0.011	0.142	-0.005	-0.023	0.159	0.213*
PSI Dys. Interaction	0.167	0.166	0.009	-0.051	0.045	0.151	-0.215*	0.097	0.160	-0.045
PSI Difficult Child	0.087	0.333**	0.016	-0.024	-0.029	0.073	0.071	0.142	0.143	0.007
PDR Daily Stress	-0.113	0.295*	0.250*	0.039	0.044	0.000	0.258*	-0.011	-0.031	0.256
PDR Daily Hassle	-0.001	0.151	0.048	-0.103	-0.027	-0.059	0.157	0.071	0.236+	0.256*
PDR Negative Mood	0.003	0.128	-0.007	-0.024	-0.131	-0.021	0.095	0.057	0.158	0.319**
PDR Positive Mood	0.012	-0.073	-0.113	-0.079	-0.095	-0.197	-0.102	-0.083	0.051	0.002
CES-D	0.164	0.083	-0.113	-0.070	-0.094	-0.117	-0.014	-0.117	0.275**	0.125
GAD-7	0.040	0.038	-0.008	-0.003	-0.175	-0.009	0.016	0.092	0.414**	0.223*
N-back % 2-back	-0.164	-0.161	-0.280*	0.009	-0.328**	-0.264*	-0.361**	0.197	0.089	0.284*
N-back % 3-back	-0.184	-0.209+	-0.011	0.034	-0.188	-0.244*	-0.370**	0.103	-0.135	0.253*
N-back % Total	-0.228+	-0.180	-0.125	-0.051	-0.351**	-0.277*	0.443**	0.196	0.068	0.281*
N-back Hits/FA	-0.276*	-0.152	-0.071	0.087	-0.332**	0.009	0.156	0.152	0.065	0.300*
WCST Total #	0.052	-0.160	-0.041	0.033	0.177	0.234*	-0.299**	0.024	-0.165	-0.334**
WCST % Accuracy	0.122	0.111	0.086	0.052	-0.195+	-0.270**	0.284**	-0.099	0.064	0.242*
WCST % Perseverative	-0.109	0.057	0.1448	0.002	0.005	-0.072	0.053	-0.054	0.023	0.251*
WCST % Conceptual	-0.125	0.128	0.113	0.005	-0.187+	-0.274**	0.251*	-0.041	0.086	0.300*

+ p <0.09, * p < 0.05, **p < 0.001

Correlations between Demographic Variables and Child Measures

Parental age was negatively correlated with child internalizing behaviors (r = -0.249, p = 0.048) and positively correlated with the ESCS Social Interaction subscales (Initiating Social Interaction: r = 0.304, p = 0.012; Responding to Social Interaction: r = 0.322, p = 0.008). Child age was also associated with a number of domains measured by the Early Social Communication Scales; older children tended to have higher scores on subscales that measured Joint Attention (Initiating Joint Attention: r = 0.269, p = 0.028; Responding to Joint Attention: r = 0.612, p <0.001) and Social Interaction (Initiating Social Interaction: r = 0.360, p = 0.003; Responding to Social Interaction: r = 0.377, p = 0.002). Child age was also marginally associated with children's response to behavior requests (r = 0.222, p = 0.070) and child externalizing behaviors (r = 0.232, p = 0.065). Child sex was significantly associated with child proficiency in responding to joint attention (t(65) = -2.000, p = 0.050), with girls responding more consistently bids for joint attention from parents (M = 7.21 , SD = 3.321) than boys (M = 5.58, SD = 3.287).

Group condition was significantly associated with two domains of child language: PLS Expressive Communication and ESCS Responding to Social Interaction. At baseline, children in the FIND group scored significantly lower on the PLS Expressive Language measure (M = 90.09, SD = 22.24) than children in the Control group (M = 100.53, SD = 13.34; t(61) = 2.232, p = 0.029). Similarly, children in the FIND group (M = 5.33, SD = 2.15) responded to significantly fewer social interaction bids from caregivers than children in the Control group (M = 6.54, SD = 2.34; t(65) = 2.153, p = 0.035). The association between group condition and PLS Total Language was marginally significant (t(60) = 1.940, p = 0.057), reflecting the same pattern of lower language scores for children in the FIND group (M = 95.84, SD = 13.57) compared to Control (M = 102.07, SD = 11.53).

Language of administration was significantly associated with several child outcomes: total problematic behaviors, externalizing behaviors, expressive communication, and initiating joint attention. Families who completed study visits in English reported higher numbers of problematic behaviors (M = 18.94, SD = 11.49) compared to those who completed study visits in Spanish (M = 11.74, SD = 12.13; t(68) = 2.025, p = 0.047). Similarly, English-speaking families endorsed higher percentages of total externalizing behaviors (M = 34.94%, SD = 20.57) than Spanish-speaking families (M = 22.82%, SD = 15.48; t(62) = 2.650, p = 0.010). Children from English-speaking families scored significantly higher on the PLS Expressive Language measure at baseline (M = 102.43; SD = 16.70) compared to children from Spanish-speaking families (M = 90.83, SD = 19.34; t(61) = 2.407, p = 0.019). The opposite was true for Initiating Joint Attention: children whose caregivers completed study visits in Spanish initiated bids for joint attention more frequently (M = 15.05, SD = 6.914) than children whose parents chose English (M = 11.44, SD = 7.21; t(65) = -2.034, p = 0.023). Finally, there was a marginally significant association between administration language and children's PLS Total Language scores, with children of English-speakers (M = 102.48, SD = 15.53) performing better than children of Spanish-speakers (M = 96.72, SD = 10.71; t(60) = 1.725, p = 0.090).

Caregiver ethnicity showed similar patterns to language in terms of its associations with child outcome variables: Caregivers who identified as Latino/a endorsed fewer child externalizing behaviors (M = 24.23%, SD = 15.43) compared to non-Latino/a caregivers (M = 37.82%, SD = 24.19; t(62) = 2.504, p =0.015). There were also marginally significant associations between caregiver ethnicity and Expressive Communication (t(61) = 1.967, p = 0.054) and Total Language (t(60) = 1.898, p = 0.062), where children with Latino/a caregivers had lower mean language scores (PLS EC: M = 92.70, SD = 19.70; PLS Total: M = 97.29 SD =

11.83) than their non-Latino/a counterparts (PLS EC: $M = 104.15$, SD = 14.05; PLS Total: $M =$ 104.77 SD = 15.45).

Parent education was positively associated with child externalizing behaviors, measured both through PDR Total Behaviors ($r = 0.270$, $p = 0.025$) and the BITSEA/CBCL Externalizing composite score ($r = 0.292$, $p = 0.020$). There was a marginally significant positive association between parent ACEs and children's tendencies to initiate social interaction ($r = 0.231$, $p = 0.060$). There was also a significant positive correlation between child adverse life experiences and child externalizing behaviors measured by the BITSEA/CBCL composite score ($r = 0.319$, $p = 0.011$).

Table 11

Correlations between Demographic Variables and Baseline Child Measures

Variable	Parent Age	Child Age	Child Sex	Group Condition	Language (Admin.)	Ethnicity	Parent Education	Household Income	Parent ACES	Child Adverse Events
PDR Total Behaviors	0.148	-0.075	-0.055	0.076	-0.239*	-0.136	0.270*	0.143	0.002	0.134
BITSEA/CBCL Internalizing	-0.249*	-0.076	0.105	0.104	0.039	-0.078	0.097	0.192	-0.090	0.052
BITSEA/CBCL Externalizing	-0.005	0.232	0.118	0.093	-0.319*	-0.303*	0.292*	0.183	0.223	0.319*
PLS Expressive Communication	0.016	0.139	-0.096	-0.275*	-0.294*	-0.244+	0.105	-0.005	0.116	-0.179
PLS Auditory Comprehension	-0.105	0.055	0.076	-0.111	-0.037	-0.118	-0.013	0.082	0.068	-0.083
PLS Total Language	-0.077	0.104	0.014	-0.243	-0.217+	-0.238+	0.050	0.031	0.085	-0.151
ESCS Init. Joint Attention	0.179	0.269*	-0.023	-0.062	0.245*	0.057	0.016	-0.076	-0.135	-0.134
ESCS Resp. Joint Attention	0.240	0.612**	0.241*	0.082	-0.018	0.003	-0.069	-0.127	0.126	0.061
ESCS Init. Behavior Request	0.125	0.054	0.125	0.094	-0.043	-0.069	0.082	0.161	-0.061	0.018
ESCS Resp. Behavior Request	-0.029	0.222+	0.017	0.071	-0.189	-0.147	0.061	-0.023	0.170	0.055
ESCS Init. Social Interaction	0.304*	0.360**	0.074	-0.087	-0.191	-0.081	0.070	-0.050	0.231	0.136
ESCS Resp. Social Interaction	0.322**	0.377**	0.159	-0.258*	0.055	0.105	0.098	0.216	-0.059	-0.065

+ p <0.09, * p < 0.05, **p < 0.001

Correlations between Demographic Variables and Dyadic Measures

Child age was significantly correlated with several key variables measured by the CT measure: Child Utterances ($r = 0.172$, $p < 0.001$), mean length of Child Utterances ($r = 0.302$, p

< 0.001), and Conversational Turns ($r = 0.706$, $p < 0.001$). Parental education was positively associated with caregiver Following behaviors ($r = 0.257$, $p = 0.022$) and negatively correlated with caregiver Leading ($r = -0.223$, $p = 0.039$). Parent ACEs had a marginally significant negative correlation with the number of Parent Utterances during play ($r = 0.193$, $p = 0.070$) and child adversity exposure was negatively correlated with Parent Utterances during dyadic play ($r = -0.328$, $p = 0.002$). There were no significant associations between parent age, child sex, group condition, or household income and any of the dyadic variables measured in this study.

There were significant correlations between caregiver ethnicity, administration language, and several dyadic variables. Caregivers who completed study visits in Spanish had significantly more Parent Utterances during baseline dyadic play ($M = 107.93$ SD = 32.67) than English-speaking parents ($M = 88.21$, SD = 35.23; t(87) = -2.580, $p = 0.012$). Spanish-speaking families also engaged in significantly more Conversational Turns ($M = 63.46$, SD = 44.66) than English-speaking families at baseline ($M = 40.07$, SD = 29.38; t(87) = -2.522, $p = 0.013$). Spanish-speaking caregivers also engaged in significantly fewer Following behaviors ($M = 35.09\%$, SD = 19.76; t(80) = 2.684, $p = 0.009$) and significantly more Leading behaviors ($M = 63.72$, SD = 19.4; t(80) = -2.196, $p = 0.005$) than English-speaking caregivers (Following $M = 48.00\%$, SD = 21.88; Leading $M = 49.42\%$, SD = 23.59). These patterns held for caregiver ethnicity as well; caregivers who identified as Latino/a engaged in higher numbers of Parent Utterances during dyadic play at baseline ($M = 108.90$, SD = 32.08) than their non-Latino/a counterparts ($M = 71.35$, SD = 28.036; t(87) = -4.438, $p < 0.001$). Latino/a caregivers also had a higher number of Conversational Turns during dyadic play ($M = 60.83$, SD = 43.40) than non-Latino/a caregivers ($M = 36.06$, SD = 27.25; t(87) = -2.246, $p = 0.027$). Latino/a caregivers also engaged in more Leading behaviors ($M = 63.35\%$, SD = 19.54; t(80) = -3.989, $p < 0.001$) and fewer Following

behaviors during play (M = 35.49%, SD = 19.85; t(80 = 3.567, $p < 0.001$) compared to their non-Latino/a counterparts (Leading M = 41.08%, SD = 22.02%; Following M = 55.23%, SD = 19.87). It should be noted that there is a significant overlap between families who identified as Latino/a and families who chose to complete study visits in Spanish (r = 0.727, p = 0.001).

Table 12

Correlations between Demographic Variables and Baseline Dyadic Measures

Variable	Parent Age	Child Age	Child Sex	Group Condition	Language (Admin.)	Ethnicity	Parent Education	Household Income	Parent ACES	Child Adverse Events
Child Utterances	0.127	0.172**	0.058	-0.114	0.165	0.084	0.075	-0.014	-0.053	0.048
Dur Chi Utterances	-0.072	0.302**	-0.168	0.034	-0.104	-0.119	0.066	0.061	0.096	0.078
Parent Utterances	0.013	0.135	0.042	-0.150	0.267*	0.430**	-0.069	0.165	-0.193+	-0.328**
Dur Parent Utterances	0.072	0.075	-0.022	0.070	-0.058	-0.078	0.124	0.027	0.095	0.112
Conversational Turns	0.061	0.706**	0.082	-0.126	0.261*	0.234*	0.023	0.037	-0.080	-0.047
Following	-0.012	0.107	-0.068	0.058	-0.287**	-0.370**	0.390**	0.151	-0.060	0.032
Leading	-0.019	-0.100	0.081	-0.069	0.312**	0.348**	-0.233*	-0.151	0.054	-0.064
Other	0.193	0.036	-0.069	0.090	0.002	0.170	-0.268	-0.090	0.035	0.246

+ p <0.09, * $p < 0.05$, **$p < 0.001$

Results for Hypotheses 1a - 1e

Aim 1 examines whether there were significant intervention effects seen within the two measured domains of parent-child dyadic interaction (i.e., Conversational Turns and FLO). Table 13 summarizes the results that show pre/post change for Child Utterances, Parent Utterances, Conversational Turns, and caregiver Following and Leading behaviors.

Table 13

Hypothesis 1: Intervention Effects on Key Dyadic Interaction Variables

Variable	β	B	SE	R^2	*p*	95% CI
Child Utterances	0.201	0.003	6.315	0.253	0.975	[-12.36, 12.77]
Parent Utterances	0.010	0.658	6.790	0.252	0.923	[-12.86, 14.18]
Conversational Turns	-0.031	-2.393	7.595	0.248	0.753	[-17.51, 12.72]
Following Behaviors	0.200	0.099	0.055	0.298	0.076+	[-0.01, 0.21]
Leading Behaviors	-0.185	-0.093	0.056	0.273	0.100	[-0.20, 0.02]

+ p <0.09, * p < 0.05, **p < 0.001

Hypothesis 1a Results (Effects of the FIND Intervention on Child Utterances)

Contrary to Hypothesis 1a, there was not a significant effect of the FIND intervention on Child Utterances across the intervention period (β = 0.003, B = 0.201, SE = 6.315, p = 0.975, 95% CI = [-12.36, 12.77]). Based on baseline correlations, child age was included as a covariate in this model; the overall model accounted for 25.3% of the total variance in post-intervention child utterance scores (R^2 = 0.253). Child age was not significantly associated with post-intervention child utterance scores in this model (β = 0.125, B = 0.413, SE = 0.444, p = 0.355, 95% CI = [-0.470, 0.501]). Pre-intervention child utterance scores were a significant predictor of post-intervention Child Utterances (β = 0.407, B = 0.454, SE = 0.151, p = 0.003, 95% CI = [0.155, 0.754]), suggesting that the quantity of child speech is fairly consistent over time. Across this study period, baseline "talkativeness" was a more robust predictor of child utterance counts than child age or whether the child received the FIND intervention.

Hypothesis 1b Results (Effects of the FIND Intervention on Parent Utterances)

There was also no significant effect of the FIND intervention on Parent Utterances across the intervention period (β = 0.010, B = 0.658, SE = 6.790, R^2 = 0.252, p = 0.923, 95%, CI = [-12.86, 14.18]). Caregiver ethnicity, home language, and child adversity exposure were

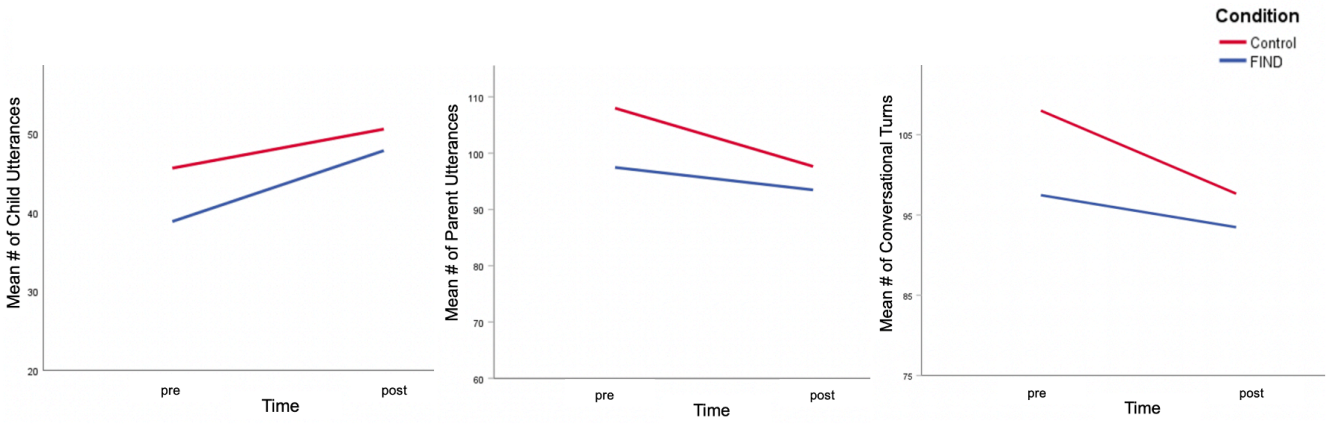
statistically indicated as covariates in this model and were included; none were significant predictors of post-intervention Parent Utterances ($p > 0.1$). As with Child Utterances, baseline Parent Utterance totals were a significant predictor of post-intervention Parent Utterance scores (β = 0.381, B = 0.354, SE = 0.105, p =0.001, 95% CI = [0.14, 0.56]), suggesting some consistency to parent talkativeness that exceeds any effect of receiving the FIND intervention.

Hypothesis 1c Results (Effects of the FIND Intervention on Conversational Turns)

Contrary to Hypothesis 1c, the FIND intervention did not have a significant effect on Conversational Turns (β = -0.031, B = -2.393, SE = 67.595, p = 0.753, 95% CI = [-17.51, 12.72]). Caregiver ethnicity, home language, and child age were included as covariates in this model, which explained 24.8% of variance in baseline Conversational Turns (R^2 = 0.248). Baseline Conversational Turns scores were a significant predictor of post-intervention Conversational Turns (β = 0.378 B = 0.350, SE = 0.130, p = 0.009, 95% CI = [0.09, 0.61]); all other covariates were not significant predictors of change in post-intervention Conversational Turns ($p > 0.3$).

Figure 7

Pre- and Post-Intervention Change in CT Variables by Intervention Condition

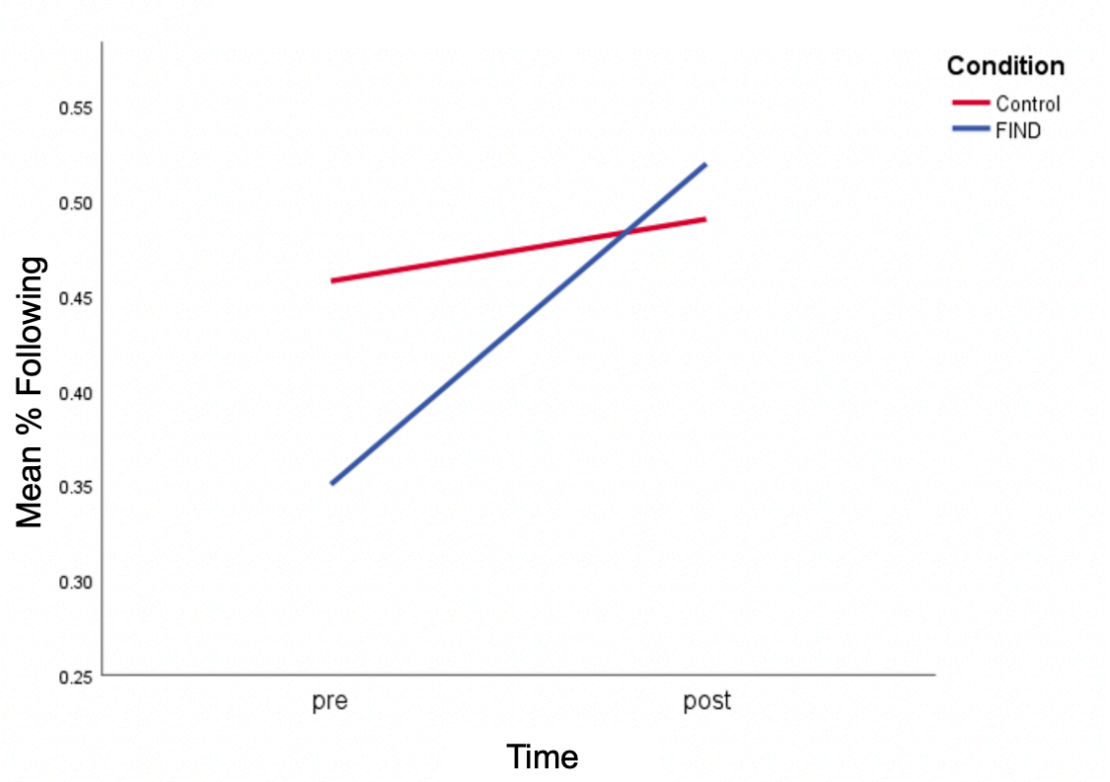


Hypothesis 1d Results (Effects of the FIND Intervention on Following Behaviors)

There was a marginally significant effect of the FIND intervention on caregiver Following behaviors (β = 0.200, B = 0.099, SE = 0.055, p = 0.076, 95% CI = [-0.01, 0.21]): caregivers in the FIND condition saw a marginally significant increase in Following across the intervention period compared to the Control group. This model accounted for 29.8% of the variance in baseline Following scores (R^2 = 0.298) and included pre-intervention Following scores, caregiver ethnicity, home language, and caregiver education as covariates. Pre-intervention Following scores were a significant predictor of post-intervention Following behaviors (β = 0.509, B = 0.578, SE = 0.139, p <0.001, 95% CI = [0.30, 0.86]), suggesting stability in this construct over time in addition to its susceptibility to intervention-related change. None of the other included covariates were significant (p > 0.6).

Figure 8

Pre- and Post-Intervention Change in Following Behavior by Intervention Condition

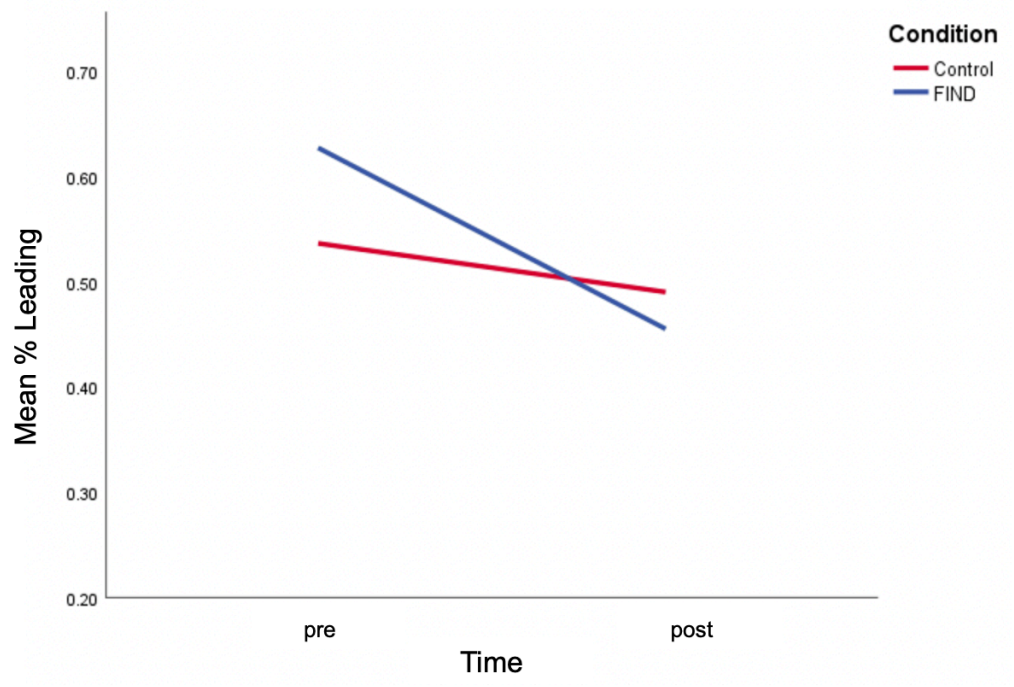


Hypothesis 1e Results (Effects of the FIND Intervention on Leading Behaviors)

Initial ANOVA analyses of pre/post intervention change suggested that caregivers in the FIND group significantly decreased their Leading behaviors across the intervention period (see Table 9). Despite these initial tests, the effect of group condition on caregiver Leading behaviors was not significant after accounting for all relevant covariates (β =-0.185, B = -0.093, SE = 0.056, p = 0.100, 95% CI = [-0.20, 0.02]). Pre-intervention Leading, caregiver ethnicity, home language, and caregiver education were included as covariates in this model, which explained 27.3% (R^2 = 0.273) of variance in post-intervention Leading scores. As with caregiver Following behaviors, baseline Leading scores were a significant predictor of post-intervention scores, suggesting some stability in the overall test re-test reliability of the FLO measure (β = 0.447, B = 0.536, SE = 0.141, p < 0.001, 95% CI = [0.26, 0.82]); no other included covariates were significant predictors (p > 0.4).

Figure 9

Pre- and Post-Intervention Change in Leading Behavior by Intervention Condition



Results for Hypotheses 2a - 2d

Aim 2 examines the extent to which dyadic interaction variables are associated with child and caregiver outcomes at the post-intervention timepoint. This aim considers *Path B* of the proposed mediation model (See Figure 1) to evaluate whether dyadic interaction is a significant predictor of other caregiver or child outcomes. Multiple linear regressions were run using post-intervention dyadic interaction scores (e.g., Conversational Turns, Following behaviors) to predict post-intervention child and caregiver outcomes (e.g., child expressive communication scores, caregiver self-efficacy). Condition (i.e., FIND or Control) was included as a covariate in all analyses to control for intervention effects on outcome variables; other covariates were included when statistically indicated based on baseline correlations between key measures and demographic variables (Tables 10, 11, 12).

Hypothesis 2a Results (Child Utterances and Child Language)

Child Utterances were not significantly associated with any post-intervention child language outcomes ($p > 0.1$; see Table 14). Across all models tested for Hypothesis 2a, the total model fit was poor (R^2 values range: 0.034 - 0.276), suggesting that neither Child Utterances nor the included covariates significantly predicted post-intervention child language scores. Child age was a significant predictor for PLS Auditory Comprehension, PLS Total Language, ESCS Joint Attention (Initiating and Responding) scores, ESCS Responding to Behavior Request, and ESCS Responding to Social Interaction ($p < 0.05$). Intervention condition (FIND vs. Control) was a significant predictor of PLS Auditory Comprehension ($p = 0.040$) and ESCS Initiating Joint Attention ($p = 0.018$).

Table 14

Hypothesis 2a: Effects of Child Utterances on Child Language

Variable	β	B	SE	R^2	*p*	95% CI
PLS Expressive Communication	0.060	0.030	0.066	0.075	0.645	[-0.10, 0.16]
PLS Auditory Comprehension	0.019	0.009	0.060	0.130	0.877	[-0.11, 0.13]
PLS Total Language	0.039	0.019	0.062	0.097	0.759	[-0.10, 0.14]
ESCS Init. Joint Attention	-0.028	-0.006	0.027	0.276	0.819	[-0.06, 0.05]
ESCS Resp. Joint Attention	-0.181	-0.019	0.013	0.227	0.154	[-0.05, 0.01]
ESCS Init. Behavior Request	-0.082	-0.002	0.029	0.009	0.935	[-0.06, 0.06]
ESCS Resp. Behavior Request	0.018	0.003	0.019	0.078	0.895	[-0.04, 0.04]
ESCS Init. Social Interaction	0.089	0.003	0.005	0.034	0.532	[-0.01, 0.01]
ESCS Resp. Social Interaction	-01.29	-0.010	0.010	0.201	0.319	[-0.03, 0.10]

+ p <0.09, * p < 0.05, **p < 0.001

Hypothesis 2b Results (Parent Utterances and Child Language)

The number of Parent Utterances during dyadic play did not significantly predict any post-intervention child language outcomes (p > 0.2; see Table 15). The total model fit was poor for most models tested for Hypothesis 2a (R^2 values range: 0.008 - 0.316), suggesting that neither Parent Utterances nor the included covariates significantly predicted post-intervention child language scores. Child age was a significant predictor for ESCS Joint Attention (Initiating and Responding) scores and ESCS Responding to Social Interaction (p < 0.02); home language was a significant predictor of ESCS Initiating Joint Attention (p = 0.029) and ESCS Initiating Social Interaction (p = 0.008), group condition was a significant predictor of ESCS Initiating Joint Interaction (p = 0.016), and caregiver ethnicity was a significant predictor of ESCS Initiating Social Interaction (p = 0.004).

Table 15

Hypothesis 2b: Effects of Parent Utterances on Child Language

Variable	β	B	SE	R^2	*p*	95% CI
PLS Expressive Communication	0.048	0.024	0.068	0.008	0.722	[-0.11, 0.16]
PLS Auditory Comprehension	0.042	0.020	0.065	0.038	0.753	[-0.11, 0.15]
PLS Total Language	0.045	0.022	0.045	0.013	0.738	[-0.11, 0.15]
ESCS Init. Joint Attention	0.105	0.023	0.105	0.316	0.414	[-0.03, 0.08]
ESCS Resp. Joint Attention	0.132	0.014	0.014	0.260	0.327	[-0.01, 0.04]
ESCS Init. Behavior Request	0.167	0.034	0.030	0.049	0.259	[-0.03, 0.09]
ESCS Resp. Behavior Request	0.034	0.005	0.020	0.044	0.819	[-0.04, 0.04]
ESCS Init. Social Interaction	-0.158	-0.005	0.004	0.194	0.264	[-0.01, 0.00]
ESCS Resp. Social Interaction	0.017	0.001	0.010	0.234	0.904	[-0.02, 0.02]

+ p <0.09, * p < 0.05, **p < 0.001

Hypothesis 2c Results (Conversational Turns and Child Language)

The number of Conversational Turns during dyadic play did not significantly predict any post-intervention child language outcomes (p > 0.2; see Table 16). As with the models from Hypotheses 1a and 1b, most models using Conversational Turns as a predictor did not explain much of the variance in post-intervention child language scores (R^2 values range: 0.042 - 0.302). As in these previous models, child age was a significant predictor for several language outcomes: PLS Auditory Comprehension, PLS Total Language, ESCS Joint Attention (Initiating and Responding) scores and ESCS Responding to Social Interaction (p < 0.03); home language was a significant predictor of ESCS Initiating Joint Attention (p = 0.031) and ESCS Initiating Social Interaction (p = 0.007). Caregiver ethnicity was also a significant predictor of ESCS Initiating Social Interaction (p = 0.006), and group condition was a significant predictor of PLS Auditory Comprehension scores (p = 0.034).

Table 16

Hypothesis 2c: Effects of Conversational Turns on Child Language

Variable	β	B	SE	R^2	*p*	95% CI
PLS Expressive Communication	0.099	0.041	0.054	0.081	0.452	[-0.07, 0.15]
PLS Auditory Comprehension	0.153	0.061	0.050	0.151	0.229	[-0.04, 0.16]
PLS Total Language	0.134	0.053	0.051	0.115	0.302	[-0.05, 0.16]
ESCS Init. Joint Attention	-0.056	-0.011	0.023	0.302	0.639	[-0.06, 0.04]
ESCS Resp. Joint Attention	-0.082	-0.008	0.011	0.252	0.514	[-0.03, 0.02]
ESCS Init. Behavior Request	-0.107	-0.019	0.025	0.042	0.447	[-0.07, 0.03]
ESCS Resp. Behavior Request	0.044	0.005	0.017	0.101	0.746	[-0.03, 0.04]
ESCS Init. Social Interaction	0.084	0.002	0.004	0.177	0.525	[-0.01, 0.01]
ESCS Resp. Social Interaction	-0.064	-0.004	0.008	0.207	0.621	[-0.02, 0.01]

+ p <0.09, * p < 0.05, **p < 0.001

Hypothesis 2d Results (Caregiver Following and Caregiver Self-Efficacy)

Caregiving Following behaviors were significantly associated with post-intervention SEPTI Teaching scores (β = 0.293, B = 6.611, SE = 2.759, p = 0.019, 95% CI = [1.10, 12.23]). This model controlled for condition, caregiver ethnicity, home language, and caregiver education and accounted for 18.8% of the variance in caregivers' post-intervention teaching self-efficacy scores (R^2 = 0.188). Following behaviors were not significantly associated with any other domains of caregiver self-efficacy or caregivers' sense of competence, although the relationship with SEPTI Nurturance was marginally significant (β = 0.234, B = 3.663, SE = 2.34, p = 0.078, 95% CI = [-0.42, 7.75]). Overall, models run to test Hypothesis 2d did not account for a significant proportion of variance in caregiver self-efficacy scores (range in R^2 = 0.034 - 0.188). None of the included covariates in these models were significant predictors of post-intervention caregiver self-efficacy scores (p > 0.1).

Table 17

Hypothesis 2d: Effects of Caregiver Following on Self-Efficacy

Variable	β	B	SE	R^2	*p*	95% CI
PSOC	0.106	1.988	2.352	0.106	0.401	[-2.71, 6.68]
SEPTI Total	0.168	12.062	9.458	0.090	0.207	[-6.84, 30.96]
SEPTI Nurturance	0.234	3.663	2.045	0.111	0.078+	[-0.42, 7.75]
SEPTI Teaching	0.293	6.611	2.759	0.188	0.019*	[1.10, 12.12]
SEPTI Instrumental Care/Routines	-0.038	-1.071	3.841	0.034	0.781	[-8.75, 6.61]
SEPTI Limit Setting	0.097	2.800	3.889	0.052	0.474	[-4.97, 10.57]

+ p <0.09, * p < 0.05, **p < 0.001

Hypothesis 2e Results (Caregiver Following and Caregiver Stress)

Caregiver Following was significantly associated with decreased overall caregiver stress levels (PSI Total Stress) at the post-intervention time point (β = -0.281, B = -25.500, SE = 11.510, p = 0.030, 95% CI = [-48.47, -2.53]). This model controlled for condition, caregiver ethnicity, caregiver education, child age, and home language and accounted for 9.6% of variance in caregivers' stress scores (R^2 = 0.096). Following behaviors also predicted lower Dysfunctional Interaction scores at post-intervention (β = -0.389, B = -12.447, SE = 3.868, p = 0.002, 95% CI = [-20.17, -4.73]). This model controlled for caregiver ethnicity, caregiver education, and home language and accounted for 15.2% of the variance in post-intervention Dysfunctional Interaction scores (R^2 = 0.152). Following behaviors were not significantly associated with any other caregiver measures of stress at the post-intervention time point (i.e., PSI Parental Distress, PSI Difficult Child, or daily reports of stress or mood from the PDR, p > 0.1). Most included covariates were not significant predictors of caregiver stress outcomes in these tested models; caregiver ethnicity (p = 0.042) and child age (p = 0.018) were significant predictors of Parental Distress measured by the PSI.

Table 18

Hypothesis 2e: Effects of Caregiver Following on Caregiver Stress

Variable	β	B	SE	R^2	*p*	95% CI
PSI Total Stress	-0.281	-25.500	11.510	0.096	0.030*	[-48.47, -2.53]
PSI Parental Distress	-0.176	-6.672	5.055	0.071	0.192	[-16.77, 3.43]
PSI Dysfunctional Interaction	-0.389	-12.447	3.868	0.152	0.002**	[-20.17, -4.73]
PSI Difficult Child	-0.136	-4.559	4.263	0.093	0.289	[-13.07, 3.95]
PDR Daily Stress	0.019	0.093	0.709	0.141	0.896	[-1.33, 1.50]
PDR Daily Hassle	0.100	1.913	2.465	0.247	0.441	[-3.04, 6.87]
PDR Negative Mood	-0.128	-2.548	2.882	0.053	0.381	[-8.34, 3.24]
PDR Positive Mood	0.024	0.676	4.105	0.005	0.870	[-7.56, 8.91]

+ p <0.09, * p < 0.05, **p < 0.001

Hypothesis 2f Results (Caregiver Following and Caregiver Mental Health)

Following behaviors were not significantly associated with either caregiver depression (measured via the CES-D) or caregiver anxiety (measured via the GAD-7) at the post-intervention time point (p > 0.3). Condition, caregiver ACEs, ethnicity, caregiver education, and home language were included as covariates in these two models; caregiver education was a significant predictor of both anxiety (p < 0.001) and depression (p = 0.033). All other covariates were non-significant (p > 0.1). These two models had an overall poor fit (R^2 values range: 0.110 - 0.259), and the majority of variance in caregiver mental health was explained by covariates and not dyadic Following behavior.

Table 19

Hypothesis 2f: Effects of Caregiver Following on Caregiver Mental Health

Variable	β	B	SE	R^2	*p*	95% CI
CESD Depression	0.034	1.082	3.984	0.110	0.787	[-6.87, 9.03]
GAD7 Anxiety	-0.104	-1.604	1.781	0.259	0.371	[-5.16, 1.95]

+ p <0.09, * p < 0.05, **p < 0.001

Hypothesis 2g Results (Caregiver Following and Caregiver Executive Function)

Following behaviors significantly predicted post-intervention success on the 2-back condition of the N-back task (β = 0.344, B = 12.119, SE = 3.854, p = 0.003, 95% CI = [4.39, 19.85). This model included condition, caregiver language, caregiver education, ethnicity, and child adversity scores as covariates and explained 49.2% of the variance in N-back scores (R^2 = 0.492). Notably, several other covariates in this model were significant predictors of performance on the N-Back: child adversity exposure (p = 0.003), home language (p = 0.047) and caregiver education (p = 0.020) were significant predictors of 2-back performance. Caregiver Following did not significantly predict any other measures of caregiver executive function (p > 0.05); Following was a marginally significant predictor of overall N-back accuracy (β = 0.214, B = 5.459, SE = 3.148, p = 0.089, 95% CI = [-0.85, 11.77]).

Across models predicting caregiver executive functioning, several covariates were consistently significant predictors of scores. Caregiver education was a significant or marginally significant predictor in all models predicting N-back performance (p-value range: 0.004 - 0.079). Child adversity exposure (and by proxy, family exposure to stressful experiences) significantly predicted N-back accuracy at 2-back, 3-back, and overall (p-value range: 0.003 - 0.025).

Table 20

Hypothesis 2g: Effects of Caregiver Following on Caregiver Executive Function

Variable	β	B	SE	R^2	*p*	95% CI
N-Back % Accuracy 2-Back	0.355	12.119	3.854	0.492	0.003**	[4.39, 19.85]
N-back % Accuracy 3-Back	0.190	6.084	4.255	0.264	0.158	[-2.45, 14.61]
N-back % Accuracy Total	0.214	5.459	3.148	0.368	0.089+	[-0.85, 11.77]
N-Back Ratio: Hits/False Alarms	0.080	2.384	3.939	0.286	0.548	[-5.52, 10.28]
WCST Total Completed Trials	-0.146	-19.691	15.892	0.204	0.220	[-51.40, 12.02]
WCST % Accuracy Total	0.061	0.037	0.074	0.137	0.623	[-0.11, 0.19]
WCST % Perseverative Response	-0.039	-0.011	0.033	0.122	0.752	[-0.08, 0.06]
WCST % Conceptual Response	0.037	0.032	0.104	0.145	0.762	[-0.18, 0.24]

+ p <0.09, * p < 0.05, **p < 0.001

Hypothesis 2h Results (Caregiver Following and Child Internalizing/Externalizing Behaviors)

Following behaviors were not significantly associated with either child internalizing behaviors (measured via a composite score of BITSEA and CBCL internalizing questions) or child externalizing behaviors (measured via a BITSEA/CBCL composite and the PDR) at the post-intervention timepoint ($p > 0.05$). Of the covariates included in these models, only caregiver ethnicity was a significant predictor of BITSEA/CBCL externalizing behaviors ($p = 0.042$). Caregiver ethnicity ($p = 0.045$) and caregiver education ($p =0.037$) were significant predictors of BITSEA/CBCL internalizing behaviors, and condition was a significant predictor of PDR Total Behaviors ($p = 0.049$). As with several of the other models run to test hypotheses in Aim 2, overall model fit was poor (R^2 values range: 0.123 - 0.149).

Table 21

Hypothesis 2h: Effects of Caregiver Following on Child Internalizing/Externalizing Behaviors

Variable	β	B	SE	R^2	*p*	95% CI
BITSEA/CBCL Composite: Internalizing	0.014	0.004	0.042	0.123	0.917	[-0.08, 0.09]
BITSEA/CBCL Composite: Externalizing	-0.153	-0.112	0.097	0.149	0.252	[-0.31, 0.08]
PDR Total Behaviors	-0.221	-9.338	5.518	0.175	0.096	[-20.41, 1.73]

+ p <0.09, * p < 0.05, **p < 0.001

Hypothesis 2i Results (Caregiver Following and Child Language)

Caregiver Following behaviors during dyadic play did not significantly predict any post-intervention child language outcomes (p > 0.1; see Table 22). Total variance explained by these models ranged from 5.9% to 32.9%, suggesting that caregiver Following (and relevant covariates) do not fully explain post-intervention child language scores. Of the covariates included in these models, child age was a significant predictor for ESCS Joint Attention scores (Initiating, p = 0.002; and Responding, p < 0.001) and ESCS Responding to Social Interaction (p = 0.006), home language was a significant predictor of ESCS Initiating Joint Attention (p = 0.022) and ESCS Initiating Social Interaction (p = 0.004), caregiver education significantly predicted ESCS Responding to Joint Attention (p = 0.005), and caregiver ethnicity was a significant predictor of ESCS Initiating Social Interaction (p = 0.004). Condition (i.e., FIND vs. Control group) was a significant predictor of ESCS Initiating Joint Attention (p = 0.021) and ESCS Responding to Social Interaction (p = 0.036). In general, models with higher R^2 values also tended to have significant covariate predictors, highlighting the non-significant relationship between Following behaviors and the tested child language outcomes.

Table 22

Hypothesis 2j: Effects of Caregiver Following on Child Language

Variable	β	B	SE	R^2	*p*	95% CI
PLS Expressive Communication	0.172	11.290	8.936	0.059	0.211	[-6.60, 29.18]
PLS Auditory Comprehension	0.026	1.604	8.477	0.062	0.851	[-15.34 18.57]
PLS Total Language	0.116	7.247	8.557	0.043	0.401	[-9.88, 24.38]
ESCS Init. Joint Attention	-0.037	-1.074	3.906	0.329	0.785	[-8.92, 6.77]
ESCS Resp. Joint Attention	0.015	0.210	1.793	0.403	0.907	[-3.40, 3.82]
ESCS Init. Behavior Request	-0.227	-6.234	4.174	0.114	0.142	[-14.62, 2.15]
ESCS Resp. Behavior Request	0.100	1.890	2.959	0.060	0.526	[-4.05, 7.83]
ESCS Init. Social Interaction	0.186	0.806	0.634	0.235	0.210	[-0.47, 2.08]
ESCS Resp. Social Interaction	-0.045	-0.426	1.321	0.295	0.748	[-3.08, 2.23]

+ p <0.09, * p < 0.05, **p < 0.001

Results for Hypotheses 3a - 3i: Mediation Models

Summary of Approach

Aim 3 evaluates the feasibility of any dyadic interaction variables (i.e., CT or FLO measures) serving as mediators of change in parent or child outcome measures. Mediation models were tested for variables where there were significant relationships seen in Path A (intervention effects on dyadic variables), Path B (associations between dyadic measures and parent or child outcomes), and Path C (direct effects of the intervention on parent or child outcomes). Rationale for running or excluding mediation analyses for Aim 3 hypotheses are discussed in more depth below.

Hypothesis 3a - 3c Results (CT measures as a Mediator for Child Language)

No significant effects were seen in Path A models for any CT measures: the FIND intervention did not have a significant effect on Child Utterances, Parent Utterances, or Conversational Turns across the intervention period (Results 1a - 1c). Further, there were no

significant effects seen in Path B models for any CT measures (Results 2a - 2c); Child Utterances, Parent Utterances, and Conversational Turns did not significantly predict post-intervention performance on any child language subscales from the PLS-5 or ESCS. There were several significant effects in child language outcomes noted across the intervention period (Path C - PLS Total Language, ESCS Initiating Joint Attention, ESCS Responding to Social Interaction and marginally significant effects on PLS Expressive Communication and PLS Auditory Comprehension). Given the lack of significant predictors in Path A and Path B, further mediation analyses were not run to evaluate the indirect effect of CT dyadic measures on change in child language scores.

Hypothesis 3d – 3j Results (FLO Following as a Mediator for Outcomes)

Path A – the effect of the FIND intervention on Following behaviors measured by FLO – was marginally significant ($p = 0.076$), suggesting a noteworthy trend in dyadic outcomes that did not reach the significance threshold ($p = 0.05$). Accordingly, no further mediation analyses were run to evaluate Following behavior as a possible mediator (Hypotheses 3d – 3i). Given varying levels of significance across Path B models in Aim 2, a further discussion of each model is offered below. Models with significant Path B and Path C analyses were included in exploratory post-hoc mediation testing to evaluate the strength of the mediation model (despite a marginally significant Path A).

Hypothesis 3d (Following as a Mediator for Caregiver Self-Efficacy).

Path B and C models were significant for SEPTI Teaching: Following behaviors significantly predicted post-intervention levels of caregiver teaching self-efficacy ($p = 0.019$; Table 17) and SEPTI Teaching scores rose significantly for caregivers in the FIND group compared to the Control group over time ($p = 0.023$; Table 7). Since the Path A model was

marginally significant, a further probe of this possible mediation effect was conducted as a post-hoc exploratory analysis (see Post-Hoc Results).

Aside from SEPTI Teaching, there were no further significant relationships seen between caregiver Following behaviors and caregiver self-efficacy scores within Aim 2d. SEPTI Teaching was also the only caregiver self-efficacy measure that showed intervention-related change between baseline and post-intervention scores (Path C; Table 7). Accordingly, no other exploratory mediation analyses were conducted to evaluate the indirect effect of caregiver Following on caregiver self-efficacy outcomes.

Hypothesis 3e Results (Following as a Mediator for Caregiver Stress).

Results 2e indicated two measures of caregiver stress that were significantly associated with caregiver Following behaviors: Total PSI and PSI Dysfunctional Interaction (Path B; Table 18). However, our baseline evaluation pre/post change from these two measures suggested that there were not main effects of the intervention on these two variables (Path C; Table 7). Accordingly, further exploratory mediation analyses were not run to evaluate the indirect effect of caregiver Following on caregiver stress scores.

Hypothesis 3f Results (Following as a Mediator for Caregiver Mental Health).

Caregiver Following behaviors did not significantly predict caregiver mental health outcomes (Path B; Table 19), nor were caregiver mental health outcomes significantly impacted by the FIND intervention in the current study (Path C; Table 7). No further exploratory mediation analyses were conducted using these variables.

Hypothesis 3g Results (Following as a Mediator for Caregiver EF).

Caregiver accuracy on the 2-back condition of the N-back task was significantly associated with caregiver Following behaviors (Path B; Table 20). However, because there were

no significant changes in 2-back accuracy scores across the intervention period (Path C; Table 7) no further mediation analyses were pursued.

Hypothesis 3h & 3i Results (Following as a Mediator for Child Outcomes).

There were no significant relationships between caregiver Following behaviors and child behaviors or child language scores (Path B; Tables 21 and 22). Although several child language variables showed significant change across the intervention period (i.e., PLS Total Language, ESCS Initiating Joint Attention, ESCS Responding to Social Interaction - see Table 8), further exploratory mediation analyses were not conducted since Path B was not significant.

Post-Hoc Analyses

In order to better understand the primary results from Aim 1, additional post-hoc analyses were conducted to probe the impact of included covariates on the impact of the overall model predicting change in caregiver Following behaviors using the FLO Coding Tool. Note that baseline repeated-measures ANOVAs that were conducted to better understand the descriptive statistics of key variables in the sample suggested that pre-post change in Following behaviors was significantly different for caregivers in the FIND compared to control groups ($p = 0.006$; Table 9). Hypothesis 1d analyses using regressions showed that these intervention-related changes were marginally significant ($p = 0.076$; Table 13) after controlling for key covariates: baseline Following scores, caregiver ethnicity, home language, and caregiver education.

Hierarchical Regressions: Predicting Post-Intervention Following Scores

First, a hierarchical regression was conducted for Hypothesis 1d to evaluate the added impact of each covariate on overall model significance. Results from this hierarchical regression are included in Table 23 and help elucidate the effect of Condition (FIND vs. Control) compared to other covariates included in these models.

Table 23

Hierarchical Regression Models Predicting Post-Intervention Following Scores

Variable	Model 1 $R^2 = 0.242$			Model 2 $R^2 = 0.285$			Model 3 $R^2 = 0.296$			Model 4 $R^2 = 0.298$		
	β	SE	p	β	SE	p	β	SE	p	β	SE	p
Baseline Following	0.492	0.118	<0.001	0.555	0.121	<0.001	0.506	0.133	<0.001	0.509	0.139	<0.001
Condition				0.217	0.053	0.045	0.201	0.054	0.066	0.200	0.055	0.076
Caregiver Ethnicity							-0.113	0.068	0.316	-0.070	0.095	0.654
Language (Admin.)										-0.072	0.075	0.624
Caregiver Education										-0.016	0.021	0.906

+ p <0.09, * p < 0.05, **p < 0.001

Notably, the only consistently significant predictor of post-intervention Following scores across models is the level of baseline Following ($p < 0.001$ across all four models). Although Condition (FIND vs. Intervention group) dips under the threshold of significance in Model 2 ($p = 0.045$), its β value remains close to 0.2 throughout Models 2, 3, and 4, suggesting that it does not explain a significant portion of the variance in post-intervention Following scores after accounting for differences between groups at baseline. The change in R^2 value between Models 1 and 2 suggests that Condition only explains an additional 4.3% of the variance in post-intervention Following scores above and beyond the predictive value of baseline Following scores. The addition of caregiver ethnicity, language of administration, and caregiver education as covariates does not significantly increase the R^2 values of the tested models, nor do their individual contributions to these models reach significance ($\beta < 0.15$ and $p > 0.3$ for Models 3 and 4). The addition of these covariates does impact the p-value associated with Condition, suggesting that placement in the FIND vs. Control group is only a moderately significant

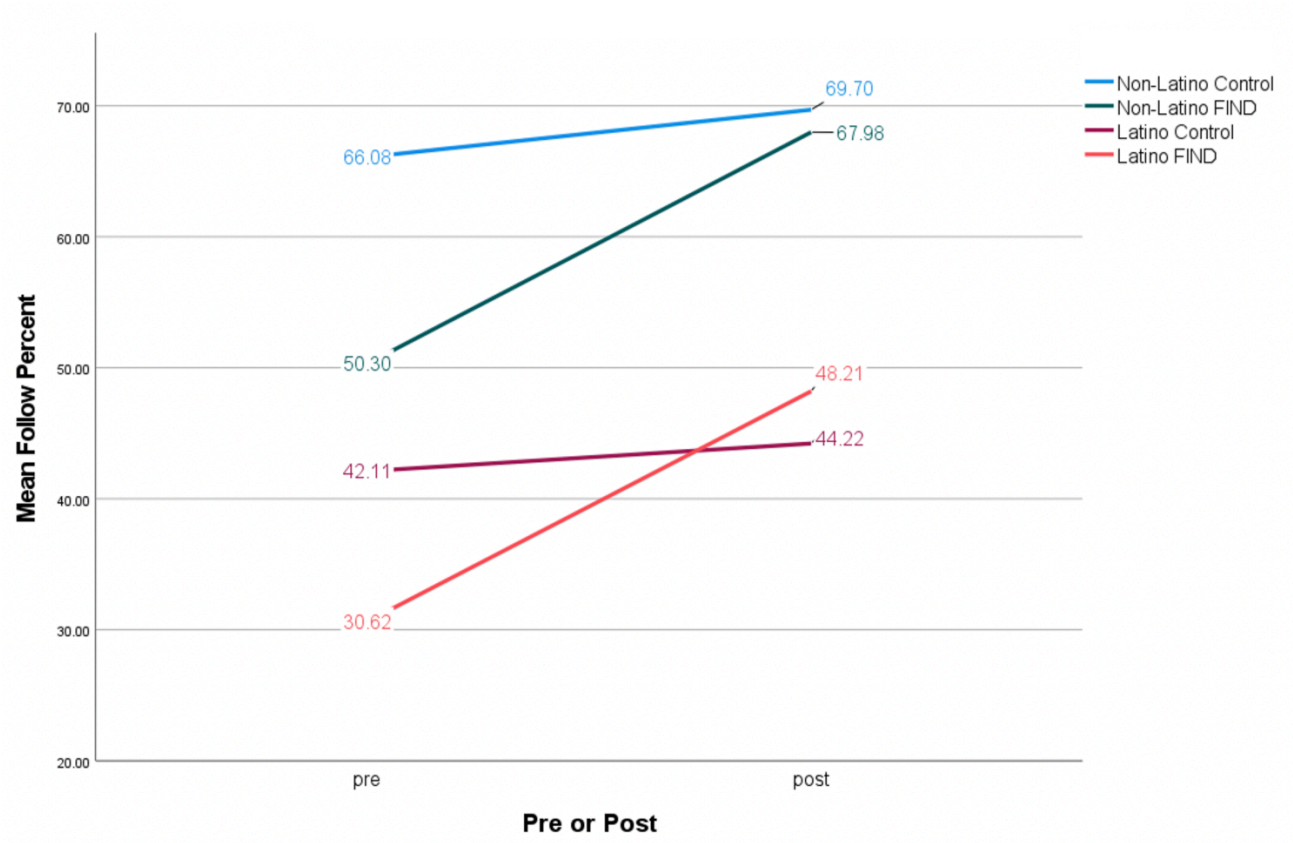
predictor of post-intervention Following scores after accounting for relevant covariates (p = 0.076).

Understanding the Relationship between Ethnicity and Caregiver Following

Second, additional analyses were conducted to better understand the differences between Latino/a and non-Latino/a groups' Following behaviors, both at baseline and in their relative change over time. Figure 10 shows the relative means for each group by Condition and ethnicity and baseline and post-intervention time points.

Figure 10

Pre-Post Change in Following Behavior by Ethnicity and Condition



Independent samples t-tests were conducted to better understand how baseline Following scores differed across groups. Group comparisons are listed in Tables 24 and 25; all tests of baseline differences were significant ($p < 0.03$), suggesting both that there were significant

differences between FIND and Control groups at baseline (t(80) = 2.295, $p = 0.025$) and significant baseline differences in Following behaviors across ethnic groups (t(80) = 3.567, $p < 0.001$).

Table 24

Relative Means for Latino and Non-Latino Groups: Following Behaviors

Ethnicity/Group	*n*	Pre-Intervention Mean (SD)	Post-Intervention Mean (SD)
Non-Latino/a Control	5	66.08% (6.48)	69.70% (22.31)
Latino/a Control	28	42.11% (21.71)	44.22% (25.76)
Non-Latino/a FIND	11	50.30% (22.14)	67.98% (21.54)
Latino/a FIND	38	30.62% (17.06)	48.21% (22.74)

Table 25

Significant Differences in Baseline Following Behaviors by Ethnicity and Condition

Ethnicity/Group	Baseline Following Mean (SD)	*t*	df	*p*
Overall Latino	35.49% (19.85)			
Overall Non-Latino	55.23% (19.87)	3.567	80	<0.001
Overall FIND	35.04% (19.87)			
Overall Control	45.74% (21.89)	2.295	80	0.024
Latino FIND	30.62% (17.05)			
Non-Latino FIND	50.30% (22.14)	3.149	47	0.003
Latino Control	42.11% (21.72)			
Non-Latino Control	66.08% (6.48)	2.421	31	0.022

+ p <0.09, * $p < 0.05$, ** $p < 0.001$

Finally, a moderation analysis was conducted to probe whether the effect of Condition (FIND vs. control) on Following behaviors is moderated by ethnicity. The interaction (Ethnicity * Condition) was not a statistically significant predictor of post-intervention Following scores (β = -0.119, B = -0.007, SE = 0.007, p = 0.300, 95% CI = [-0.02, 0.01]), suggesting that the effect of FIND on caregiver Following behaviors does not change significantly by ethnicity. This is corroborated by the largely parallel lines in Figure 10 that show similar change in Following behaviors for Latino/a and non-Latino/a caregivers within each condition.

Exploratory Mediation: SEPTI Teaching and Caregiver Following

Because Path A analyses for dyadic interaction variables did not reach the threshold of significance, no mediation analyses were conducted for Hypotheses 3a – 3i. However, the marginally significant intervention effects on FLO Following behaviors are worth further post-hoc evaluation. One tested outcome variable met the threshold for further mediation testing based on (marginally) significant results for Paths A, B, and C: caregiver self-efficacy for teaching (SEPTI Teaching).

Results for Hypothesis 2d suggested a significant relationship between FLO Following behaviors and post-intervention SEPTI Teaching (Path B: β = 0.293, B = 6.611, SE = 2.759 p = 0.019, 95% CI = [1.10, 12.12]). Baseline descriptive analyses examining intervention-based change in SEPTI Teaching scores suggested a significant Path C (Table 7). A multivariate regression confirmed the significance of the total effect (Path C) of FIND on SEPTI Teaching (β = 0.257, B = 2.771, SE = 1.173, p = 0.021, 95% CI = [0.43, 5.11]).

Accordingly, a mediation analysis was run to evaluate whether caregiver Following behaviors mediated the change in SEPTI Teaching scores across the intervention period. This model included home language, caregiver ethnicity, and caregiver education as covariates, which

were statistically indicated based on baseline correlations between these demographic variables and key outcome measures (Tables 10 and 12). The overall model was significant, ($F(5, 57) = 6.611$, $R^2 = 0.3670$, $p = 0.001$), suggesting that 36.7% of variance in post-intervention SEPTI Teaching scores are explained by group condition and caregiver Following behaviors. The model summary for Path A showed a marginally significant effect of the FIND intervention on Following behaviors over time ($t(57) = 1.803$, $p = 0.077$), which is consistent with the Path A analyses conducted in Aim 1. The model summary for Path B showed a significant effect of caregiver Following behaviors on post-intervention SEPTI Teaching scores ($t(56) = 2.573$, $p = 0.013$), which is consistent with the Path B analyses conducted in Aim 2. The direct effect (Path C') within this mediation was not significant ($t(56) = 0.849$, $p = 0.339$), suggesting that caregiver Following may play a significant role in explaining the relationship between FIND and SEPTI Teaching. The indirect path that evaluated the effect of group condition on SEPTI Teaching scores mediated by caregiver Following was also not significant (CI: [-0.10, 2.52][*]).

Table 26

Post-Hoc Mediation Analysis: SEPTI Teaching Scores Mediated by Caregiver Following

Mediation Path	B	SE	*t*	*p*	95% CI
Path A	0.099	0.055	1.803	0.077+	[-0.01, 0.21]
Path B	9.089	3.532	2.573	0.013*	[2.01, 16.16]
Path C'	1.273	1.498	0.850	0.399	[-1.73, 4.27]
Indirect Effect	0.895	0.672			[-0.10, 2.53]

+ $p < 0.09$, * $p < 0.05$, ** $p < 0.001$

[*] Note that the PROCESS model includes estimated 95% confidence intervals instead of reported *p* values to indicate overall significance of the calculated indirect effect.

CHAPTER 4: DISCUSSION

Since the development of the FIND intervention in 2016, there has been a growing body of empirical evidence supporting its positive effects on both participating caregivers and their children. Yet various efforts to measure the mechanism underlying these changes have fallen short, due in part to the complexities of developing robust video-coding methodologies that were feasible, reliable, and closely aligned with the FIND theory of change. The current study is the culmination of years of programmatic work to develop effective measurement tools that allow us to evaluate responsive caregiving in the context of dyadic play interactions. This is the first study that attempts to evaluate how the FIND intervention impacts caregivers' responsiveness and investigates whether or not changes in dyadic interaction may mediate observed changes in child and caregiver outcomes.

The current study had several important findings. Contrary to our hypotheses, Aim 1 found that there were no significant changes in the quantity of Utterances or Conversational Turns between pre- and post-intervention time points. There was a marginally significant increase in caregiver Following behaviors among families in the FIND group compared to the Control group. For Aim 2, there were no significant associations between any CT measures (e.g., Child or Parent Utterances, Conversational Turns) and any child language outcomes. Increased caregiver Following behaviors significantly predicted increased caregiver self-efficacy in teaching, decreased parental stress, decreased parent-child dysfunctional interaction, and increased caregiver accuracy in the 2-back task, which draws on working memory and sustained attention. Aim 3 considered the significance of findings from Aims 1 and 2 to evaluate whether dyadic interaction may serve as a mediator for any caregiver or child outcomes. Because no dyadic interaction variables showed statistically significant intervention effects, further

mediation analyses were not conducted. A post-hoc exploratory mediation analysis was conducted for caregiver self-efficacy in teaching, which met minimum thresholds for mediation with a marginally significant Path A effect; neither the direct or indirect effects of FIND or Following-behavior-through-FIND met the threshold for significance, although results showed promising evidence of a marginally significant mediation effect. Further interpretation of these results, including a detailed analysis of changes in key variables over time, covariates, and an interpretation of these findings in the context of prior literature is included below.

Dyadic Utterances and Conversational Turns: Key Takeaways

Analyses from Aim 1 provided one of this study's key takeaways: the FIND intervention did not significantly impact the quantity of parent or child speech or the number of verbal back-and-forth interactions during play. Means for Parent Utterances, Child Utterances, and Conversational Turns were remarkably consistent across the intervention period - with no significant differences between groups. The range in change scores across participants (e.g., -91 to 113 for Conversational Turns) suggests that the measure is sensitive to change in individuals over time, *and* that most dyads had fairly consistent speech patterns across pre- and post-intervention assessments (Figure 7; Table 9).

Aim 2 probed whether Child Utterances, Parent Utterances, and Conversational Turns were predictive of other child language outcomes at the post-intervention time point. Contrary to hypotheses – and prior validation efforts for the CT measure – none of these three measured dyadic variables were associated with any child language outcomes. It is somewhat surprising that seemingly robust prior relationships between Child Utterances, Conversational Turns, and measures of child language did not show consistent effects in the current study; further research with a larger sample size and adequate statistical power would help elucidate remaining

questions about these relationships. It is worth noting that several key measures of child language in this study were significantly associated with ethnicity and home language, including Parent Utterances and Conversational Turns. Future research with more intentionally balanced cultural groups (e.g., Latino/a vs. non-Latino/a) should look into how ethnicity, home language, dyadic language interactions, and child outcomes interact to help clarify the inconsistencies seen between this study's findings and prior validated work on Conversational Turns and dyadic talkativeness (Beiting et al., 2022; Ferjan Ramírez et al., 2020; Romeo et al., 2021).

It is also worth noting that the bulk of existing literature supporting connections between dyadic speech patterns and child language scores uses long-form recording tools (e.g., Language Environment Analysis – LENA; see Cristia et al., 2021), while this study coded language interactions across short 7-minute play sessions. While some researchers have highlighted important methodological differences between structured play and naturalistic routines (Tamis Le-Monda et al., 2017), there are no direct comparisons that consider how language counts during short, structured play sessions may compare to long-form LENA recordings. Efforts to directly compare shorter CT-coded play interactions and daylong automated language estimates via LENA audio files are currently underway (McDorman et al., 2024). The inconsistency between this study's findings and LENA-based research on dyadic speech variables highlights important gaps in our understanding of how methodology impacts our interpretation of parent-child language interactions and home language environments.

One critical limitation of the current CT coding system is its reliance on vocalizations to capture serve and return interactions. Across the FIND intervention, child "serves" include all bids for caregiver attention, including gestures, shifts in attention, vocalizations, and movement. The decision to use the CT coding system as a proxy for serve and return interactions was rooted

in prior challenges with inter-rater reliability and coding feasibility for coding measures that involved more nuanced evaluation of gestures and vocal utterances (Imhof, 2019). Recent research highlights the limitations of relying on solely verbal back and forth interaction, noting that bimodal serves (i.e., vocal-gestural combinations) elicit different responses from caregivers (van der Klis et al., 2023). These findings caution against research approaches that flatten the caregiver-child interaction landscape into the verbal-only communication, especially for young children who communicate in such rich, multidirectional ways (e.g., through speech, gestures, looking, and affect). The results from the current study reiterate the limitations associated with measuring verbal back-and-forth interactions alone and highlight the need for more nuanced serve and return coding in future evaluations of FIND.

Unpacking FIND and FLO Results: Implications for Following the Child's Lead

The Impact of FIND on Following the Child's Lead

A second key takeaway from the current study is that the FIND intervention seems to have a promising impact on caregiver Following behaviors during dyadic play. After controlling for covariates, the effect of FIND on caregiver Following was marginally significant ($p = 0.076$). Post-hoc analyses revealed a consistent increase in Following behavior for caregivers who received the FIND intervention across Latino/a and non-Latino/a groups, despite significant differences between ethnic groups at baseline. These trends - and the lack of significance of ethnicity as a moderator of intervention effects - suggest a systematic change in the way that caregivers interact with their children across the FIND intervention period that transcends pre-existing cultural differences in parenting interactions.

It should be noted that there was a significant difference in baseline Following behaviors between the FIND and Control groups; caregivers assigned to the FIND condition consistently

engaged in more Leading behaviors and fewer Following behaviors during dyadic play at baseline, a trend that remained consistent across both Latino/a and non-Latino/a subgroups (Figure 10). Although the current study controlled for baseline scores in analyses, there are several possible explanations for our current results: a) the current study results primarily represent a regression to the mean effect; b) the FIND intervention is most effective for caregivers with low base rates of Following their children's lead; or c) the FIND intervention effectively increases caregiver Following behaviors during dyadic play regardless of baseline scores. Post-hoc analyses breaking down intervention effects by ethnicity (Figure 10) provide some evidence for the third hypothesis - that FIND effectively increases caregiver Following behaviors across varying levels of baseline Following scores. However, unequal sample sizes for each group due to the current study's majority-Latino/a participants make it difficult to draw definitive conclusions about the effects of FIND on caregivers with higher baseline rates of Following behaviors during play. Future research should aim to replicate this study's findings with a larger sample that has more balanced parenting behaviors at baseline between control and intervention groups. Intervention effects on Following behaviors that persist in a study with equal baseline variances between groups would provide further evidence that the current study's findings are indicative of true intervention effects beyond regression to the mean (Barnett et al., 2005).

Post-hoc hierarchical regression analyses helped to delineate the relative impact of intervention group effects compared to other predictors in our model. Overall, baseline Following scores were consistently the strongest predictor of post-intervention Following behavior ($p<0.001$, $R^2 = 0.242$); the relative increase in R^2 after adding group condition as a predictor (+ 0.43) suggests that intervention effects explain only a small percentage of the overall

variance in post-intervention caregiver Following scores. The fact that baseline scores were such a robust predictor of post-intervention scores indicates that dyadic Leading and Following behaviors are fairly stable over time and provides further evidence that the influence of group condition was relatively less impactful.

Following and its Relationship with Caregiver and Child Outcomes

Aim 2 revealed several significant associations between post-intervention caregiver Following behaviors and other caregiver outcomes. In particular, caregivers who followed their child's lead during play more often reported lower levels of overall stress and dysfunctional interaction (measured via the PSI) and higher levels of teaching self-efficacy (measured via the SEPTI) at the post-intervention timepoint. Relationships between Following behaviors and teaching self-efficacy are consistent with prior validation efforts for the FLO Coding Tool (Table 3), while the negative association between Following and caregiver stress is a new finding in the current study. There was also a significant relationship between caregiver Following and accuracy scores on the 2-back trials of the N-back task, although the lack of consistency with other measures of working memory and attentional control across other N-back subscales (e.g., overall accuracy, 3-back accuracy) provides inconclusive evidence about the strength of this effect. Further evaluation of the association between Following behaviors and caregiver executive functioning in the context of a study with a larger sample size would help clarify the test-retest reliability of these relationships.

Relationships with child outcomes showed a similar pattern of inconsistency with prior validation efforts for FLO. Previously robust correlations between Following and child internalizing behaviors (Table 4) were not present at the post-intervention timepoint (Table 21). Similarly, previously-found associations between Following, PLS Expressive Communication,

and ESCS Initiating Social Interaction were not present at the post-intervention timepoint in the current study (Table 22). Surprisingly, Following behavior was not significantly associated with any measured child outcomes (language or behavior) at the post-intervention timepoint of this study, a marked departure from prior knowledge about the relationship between FLO behaviors and child outcomes (Imhof et al., 2023).

The fact that there was a marginally significant change in Following scores across time while many other outcome measures remained constant (Tables 7, 8, and 13) could explain why many of the expected associations tested in Aim 2 were non-significant. Validation efforts for the FLO Coding Tool used baseline scores – from the current study, among other datasets – between FLO behaviors and other outcomes to understand how caregiver tendencies during dyadic interactions were associated with other caregiver and child traits. Strong associations between caregiver Following and other key measures at one time point do not guarantee that those same associations will remain as each outcome changes differentially over time. For example, a prior validation of the FLO Coding Tool found a robust correlation between caregiver Following and child internalizing behaviors (r = 0.460; Table 4, Imhof et al., 2023). In the current study, there were marginally significant intervention effects on Following, but no significant intervention effects on child internalizing scores. Given differential changes in Following and internalizing behaviors over the 10-week intervention period, it logically follows that these two measures would no longer be significantly correlated at the post-intervention timepoint.

Given this, it is possible that caregiver Following could still serve as a mediator of change driving downstream child outcomes, although the mechanism of change (i.e., Following the child's lead) may need more than 10 weeks to make a visible impact. This study collected

data from families at pre-intervention and post-intervention timepoints; future research could help clarify whether changes to dyadic interaction mediate change in child outcomes over time by adding a follow-up visit several months after the post-intervention timepoint to re-evaluate these key relationships between measures.

Following as a Potential Mediator: Unpacking Aim 3

Although one of the current study's primary aims included evaluating measures of FLO and CT as possible mediators of change in caregiver and child outcomes, nonsignificant intervention effects across key variables precluded the majority of the proposed mediation analyses. The theory underpinning the hypotheses for this study was based in prior literature that found significant intervention effects of FIND on caregiver stress, self-efficacy, mental health, and executive function (Giuliani et al., 2019; Liu, Phu, et al., 2021; Schindler et al., 2017, 2021) along with impacts on child behaviors and child language (Imhof et al., 2023; Liu, Fisher, et al., 2021; Schindler et al., 2017, 2021). Baseline descriptive analyses for this study revealed far fewer significant intervention effects on key outcome measures than expected (Tables 7, 8, and 9). Across all expected intervention outcomes, only parent sense of competence (PSOC), teaching self-efficacy (SEPTI), overall child language (PLS Total), and two measures of the ESCS (initiating joint attention and responding to social interaction) showed significant differences between the FIND and Control groups over time. Without direct intervention effects on child and caregiver outcomes (Path C), many of the proposed mediation analyses for the current study became irrelevant.

After accounting for the results from Aim 1 and Aim 2, only caregiver self-efficacy for teaching (SEPTI Teaching) met the criteria for post-hoc exploratory mediation testing (i.e., significant or marginally significant Path A, B, and C effects). Mediation analyses using

PROCESS (Hayes, 2012) revealed consistent findings with the current study's initial Path A (Aim 1) and Path B (Aim 2) effects. The fact that the total effect (Path C) is significant while the direct effect (Path C') was not significant suggests that the mediator (i.e., caregiver Following) may explain a portion of the changes in SEPTI teaching scores across the intervention period. Notably, the indirect effect (i.e., the effect of FIND on SEPTI Teaching through caregiver Following) was also not significant, a somewhat unusual pattern of findings within mediation testing. In this case, is important to note that Path A was only *approaching* significance (p = 0.077), precluding the possibility of a fully significant mediation. This probe for mediation effects found an indirect effect that was also approaching significance (95% CI [-0.1, 2.53]), although it did not meet the threshold to be considered a significant mediator of change.

Overall, this evaluation of dyadic interaction as a possible mediator of change provides promising evidence that caregiver Following behaviors may partially mediate FIND-related changes in caregiver teaching self-efficacy. Future studies considering the overall impact of FIND on dyadic caregiver-child interactions should continue to probe this possible mediation relationship given the preliminary evidence provided by the current study.

Limitations of the Current Study

A critical limitation of the current study stems from the fact that our two groups had significantly different baseline FLO Following scores. While our results appear to indicate promising FIND-related increases in Following behaviors, it is somewhat difficult to understand the main effects of the intervention because the Control and FIND groups started in very different places. Past evaluations of FIND have found some evidence that the intervention is best suited for caregivers with the highest baseline risk factors (e.g., lower self-efficacy, lower parenting competence; Schindler et al. 2017); it may be true that FIND also works best for

families who start with lower baseline rates of Following. Although the current study provides promising evidence of FIND-related changes to Following, we should interpret these results and their wider implications with caution because of these baseline group differences. Several large-scale RCTs of FIND are currently underway that should offer opportunities to evaluate the robustness of these findings in a sample with more balanced FLO behaviors across groups.

Another critical limitation of this study is the overall sample size and prevalence of missing data. Power analyses conducted for the original grant proposed a sample size of 180 families enrolled, with at least 60 families in each group after accounting for participant attrition. 138 families were actually enrolled in the study, and attrition rates were higher than expected (33%), resulting in only 38 families in the control group and 54 families who received the full FIND intervention. Even among the 92 families who completed all study visits, poor video quality limited the number of codeable interaction films, and many families had incomplete or missing data across several measures collected during research visits (Table 5). As a result, analyses conducted for the current study are underpowered, impacting our ability to draw wider conclusions from the trends reported here.

The demographic composition of the current sample offers both strengths and limitations to the findings associated with this research. Because such a high percentage of the caregivers in this sample identified as Latino/a (80.4%), this study offers valuable insights into the efficacy of the FIND intervention with Latino/a populations. Previous study efforts have shown promising effects of FIND with Latino/a families (Schindler et al., 2017, 2021), but this is the first evidence within the context of a larger RCT that FIND promotes beneficial caregiver and child outcomes for Spanish-speaking Latino/a groups. The proportion of Latino/a families to non-Latino/a families does have limitations in the context of this study, where sample size is already a limiting

factor. Because of the small number of caregivers identifying as non-Latino/a (5 control, 11 FIND), it is difficult to draw robust conclusions from a comparison across groups. The current study highlighted several significant differences between Latino/a and non-Latino/a caregivers, particularly in the context of parenting behaviors. At baseline, Latino/a caregivers had higher levels of Parent Utterances and Conversational Turns during dyadic play; Latino/a caregivers also engaged in more Leading behaviors and fewer Following behaviors compared to their non-Latino/a counterparts. These comparisons offer rich insights into how cultural parenting values impact dyadic interaction at the microsocial level, although it is difficult to draw broad conclusions from a comparison of such uneven sample sizes. Future research should consider a more intentional design to compare Latino/a and non-Latino/a families, both across measures of microsocial dyadic interactions and in their response to the FIND intervention.

Parenting Behaviors Across Cultures: Cultural Considerations for FIND

The demographic composition of this study also warrants further discussion about the implications of our findings in the context of known differences in parenting values across cultures. The current study found systematic differences in dyadic interaction variables across Latino/a and non-Latino/a families; most notably, Latino/a caregivers tended to engage in more Leading behaviors and fewer Following behaviors during dyadic play, an observation that was supported by qualitative observations from our bilingual and bicultural coding teams during the video-coding process. This finding alone is not surprising - there are well-documented differences in parenting behaviors and values across cultures (e.g., Calzada et al., 2012; Lansford, 2022; Zayas et al., 2009), many of which emphasize Latino/a caregivers' increased directiveness during dyadic interactions (Dyer et al., 2014; Peredo et al., 2020; Tamis-LeMonda et al., 2020). The replication of these findings in the context of an intervention that targets child-

centered “Following” behaviors (i.e., the current study) yields questions about the complex interplay between definitions of “*responsive parenting”* and the ethics of working within - and in some cases, trying to change - cultural parenting norms.

More recent reviews of cross-cultural parenting practices highlight the disproportionate representation of middle-class, White families in the psychological studies that underlie our seminal theories (Cabrera, 2022; Lansford, 2022). Attachment theory and its associated definitions of “responsiveness” and “maternal sensitivity” have been critiqued for offering narrow definitions that promote specific caregiving behaviors (e.g., following the child’s lead) that are more culturally relevant to Western, urban populations (Keller, 2018). This recent emphasis on better understanding how sensitive and responsive parenting practices differ across cultures (e.g., Mesman et al., 2018) sheds light on our need to engage with deeper reflective practice as we develop, evaluate, and culturally adapt parenting interventions. Lansford’s 2022 review concludes with a series of recommendations for researchers developing parenting interventions, emphasizing the need to a) situate interventions within the sociocultural context of its participants; b) prioritize definitions of responsive parenting that promote nurturance and positive parenting practices; and c) continually adapt programs through systematic monitoring and evaluation.

Lansford’s suggestions provide a helpful framework for self-reflection about the strengths and limitations of this study’s approach to sociocultural competence. It is worth noting that many aspects of FIND - namely its strengths-focused, individualized approach - offer opportunities for microsocial adaptations and individualization to align feedback with family cultural values. The fact that caregivers are receiving strengths-focused feedback by watching videos of their own interactions with their children places FIND families at the center of the

curriculum and sets FIND apart from more rigid, didactics-focused intervention programs. These opportunities to individualize feedback through personalized video coaching underlie the broad scalability of FIND and its success in various implementations across broad geographic and situational contexts. Indeed, the fact that FIND has shown promising results in implementations across cultural contexts (e.g., Schindler et al., 2017, 2021) suggests the broader scalability and cross-cultural feasibility of its approach and theory of change.

Ultimately, findings from the current study that highlight systematic differences between Latino/a and non-Latino/a caregivers in a domain central to the FIND theory of change: following the child's lead during play. It is critical to situate these findings in the context of broader conversations about cultural competence and social justice work among minoritized, marginalized populations in the United States. Aligning with the recommendations from Lansford and other researchers emphasizing the importance of cross-cultural humility in parenting intervention work, future studies of FIND should include more intentional efforts to monitor and evaluate the intervention's impact on broader cultural constructs. Collection of qualitative data that documents caregiver reactions to the FIND curriculum, alignment with cultural parenting practices, and feedback would help FIND evolve to include more culturally responsive and inclusive definitions of sensitivity.

Implications and Conclusions

Overall, the current study offers several promising first steps in the attempt to effectively capture FIND-related changes in responsive caregiving over time. After years of programmatic work to refine and develop observational coding tools that were well-aligned to the FIND theory of change, this study is the first to document FIND-related changes in parent-child interaction. The current study also provides promising evidence that FLO Following may mediate

intervention-related changes in caregiver teaching self-efficacy. Although the mediation effects reported here are only marginally significant, these results suggest that caregiver Following behaviors may serve as a mechanism of change driving FIND-related improvements in caregiver self-efficacy. Given the noted limitations associated with the current study sample, these findings provide a strong basis for future replication work with the two larger randomized-controlled trials of FIND currently underway.

This is also the first large-scale implementation of the two observational coding tools at the center of this study's methodology. Although we did not see expected intervention-related effects on CT variables, the current study offers evidence that a) both observational coding tools are sensitive to change over short periods and b) the FLO Coding Tool can effectively detect FIND intervention effects. Results linking Following and Leading behaviors to several key caregiver and child outcomes suggest that the FLO Coding Tool effectively captures relevant components of responsive caregiving and may serve as a mediator of change in other intervention contexts. Ultimately, these findings herald a promising trajectory for the FLO Coding Tool, highlighting its potential for widespread application in assessing responsive caregiving behaviors within clinical intervention research.

REFERENCES

Abidin, R. R., Fiens, J. R., & Austin, W. G. (2006). The parenting stress index. In R. P. Archer (Ed.), *Forensic uses of clinical assessment instruments* (pp. 297-328). Lawrence Erlbaum Associates Publishers.

Achenbach, T. M. (2001). *Manual for the ASEBA school-age forms & profiles: An integrated system of multi-informant assessment.* Burlington: University of Vermont, Research Center for Children, Youth, & Families.

Ainsworth, M. D. (1963). The development of infant-mother interaction among the Ganda. In *Determinants of Infant Behavior* (pp. 67–112). Wiley.

Ainsworth, M. D. (1985). Patterns of attachment. *Clinical Psychologist, 38*(2), 27-29.

Ainsworth, M. D. S., Bell, S. M., & Stayton, D. F. (1974). Infant-mother attachment and social development: Socialization as a product of reciprocal responsiveness to signals. In *The integration of a child into a social world* (pp. 99–135). Cambridge University Press.

Ainsworth, M. D. S., Blehar, M. C., Waters, E., & Wall, S. (1978). *Patterns of attachment: A psychological study of the strange situation* (pp. xviii, 391). Lawrence Erlbaum.

Allen, B., Timmer, S. G., & Urquiza, A. J. (2014). Parent–Child Interaction Therapy as an attachment-based intervention: Theoretical rationale and pilot data with adopted children. *Children and Youth Services Review*, *47*, 334–341. https://doi.org/10.1016/j.childyouth.2014.10.009

Anderson, N. J., Graham, S. A., Prime, H., Jenkins, J. M., & Madigan, S. (2021). Linking quality and quantity of parental linguistic input to child language skills: A meta-analysis. *Child Development*, *92*(2), 484–501. https://doi.org/10.1111/cdev.13508

Artiola i Fortuny, L., & Heaton, R. K. (1996). Standard versus computerized administration of

the Wisconsin Card Sorting Test. *Clinical Neuropsychologist*, *10*(4), 419–424. https://doi.org/10.1080/13854049608406702

Asok, A., Bernard, K., Roth, T. L., Rosen, J. B., & Dozier, M. (2013). Parental responsiveness moderates the association between early-life stress and reduced telomere length. *Development and Psychopathology*, *25*(3), 577–585. https://doi.org/10.1017/S0954579413000011

Axberg, U., Hansson, K., Broberg, A. G., & Wirtberg, I. (2006). The development of a systemic school-based intervention: Marte Meo and coordination meetings. *Family Process*, *45*(3), 375–389. https://doi.org/10.1111/j.1545-5300.2006.00177.x

Bagwell-Gray, M. E., Grube, W., Mendenhall, A., Jen, S., Olaleye, O., & Sattler, P. (2023). A qualitative exploration of caregivers' experiences with the Attachment and Biobehavioral Catch-up (ABC) parenting program. *Infant Mental Health Journal*, *44*(3), 406–421. https://doi.org/10.1002/imhj.22057

Balldin, S., Fisher, P. A., & Wirtberg, I. (2018). Video feedback intervention with children: A systematic review. *Research on Social Work Practice*, *28*(6), 682–695. https://doi.org/10.1177/1049731516671809

Bandura, A. (1965). Influence of model's reinforcement contingencies on the acquisition of imitative responses. *Journal of Personality and Social Psychology*, *1*, 589–595. https://doi.org/10.1037/h0022070

Bandura, A., Ross, D., & Ross, S. A. (1963). Imitation of film-mediated agressive models. *Journal of Abnormal and Social Psychology*, *66*, 3–11. https://doi.org/10.1037/h0048687

Barnard, K. E. (1997). Influencing parent-child interactions for children at risk. In M. J. Guralnick (Ed.), *The effectiveness of early intervention* (pp. 249-268). Paul H. Brookes

Publishing.

Barnard, K. E. & Eyres, S. J. (1979). *Child health assessment part 2: The first year of life.* (Publication No. DHEW HRA 79-25). U.S. Government Printing Office.

Barnett, A. G., van der Pols, J. C., & Dobson, A. J. (2005). Regression to the mean: What it is and how to deal with it. *International Journal of Epidemiology*, *34*(1), 215–220. https://doi.org/10.1093/ije/dyh299

Beiting, M., Alper, R. M., Luo, R., & Hirsh, -Pasek Kathy. (2022). Keep the ball rolling: sustained multiturn conversational episodes are associated with child language ability. *American Journal of Speech-Language Pathology*, *31*(5), 2186–2194. https://doi.org/10.1044/2022_AJSLP-21-00333

Bernard, K., Dozier, M., Bick, J., & Gordon, M. K. (2015). Intervening to enhance cortisol regulation among children at risk for neglect: Results of a randomized clinical trial. *Development and Psychopathology*, *27*(3), 829–841. https://doi.org/10.1017/S095457941400073X

Bernard, K., Dozier, M., Bick, J., Lewis-Morrarty, E., Lindhiem, O., & Carlson, E. (2012). Enhancing attachment organization among maltreated children: Results of a randomized clinical trial. *Child Development*, *83*(2), 623–636. https://doi.org/10.1111/j.1467-8624.2011.01712.x

Bernard, K., Hostinar, C., & Dozier, M. (2015). Intervention effects on diurnal cortisol rhythms of CPS-referred infants persist into early childhood: Preschool follow-up results of a randomized clinical trial. *JAMA Pediatrics*, *169*(2), 112–119. https://doi.org/10.1001/jamapediatrics.2014.2369

Bernard, K., Lee, A. H., & Dozier, M. (2017). Effects of the ABC intervention on foster

children's receptive vocabulary: Follow-up results from a randomized clinical trial. *Child Maltreatment*, *22*(2), 174–179. https://doi.org/10.1177/1077559517691126

Bigelow, A. E., MacLean, K., Proctor, J., Myatt, T., Gillis, R., & Power, M. (2010). Maternal sensitivity throughout infancy: Continuity and relation to attachment security. *Infant Behavior and Development*, *33*(1), 50–60. https://doi.org/10.1016/j.infbeh.2009.10.009

Blaisdell, K. N., Imhof, A. M., & Fisher, P. A. (2019). Early adversity, child neglect, and stress neurobiology: From observations of impact to empirical evaluations of mechanisms. *International Journal of Developmental Neuroscience: The Official Journal of the International Society for Developmental Neuroscience*, *78*, 139–146. https://doi.org/10.1016/j.ijdevneu.2019.06.008

Bodrova, E., Leong, D. J., Hensen, R., & Henninger, M. (2000). Imaginative, child-directed play: Leading the way in development and learning. *Dimensions of Early Childhood*, *28*(4), 25–30.

Bohr, Y., Putnick, D. L., Lee, Y., & Bornstein, M. H. (2018). Evaluating caregiver sensitivity to infants: Measures matter. *Infancy: The Official Journal of the International Society on Infant Studies*, *23*(5), 730–747. https://doi.org/10.1111/infa.12248

Bornstein, M. H., Putnick, D. L., Cote, L. R., Haynes, O. M., & Suwalsky, J. T. D. (2015). Mother-infant contingent vocalizations in 11 countries. *Psychological Science*, *26*(8), 1272–1284.

Bornstein, M. H., & Tamis-LeMonda, C. S. (1989). Maternal responsiveness and cognitive development in children. *New Directions for Child Development*, *43*, 49-.

Bornstein, M. H., & Tamis-LeMonda, C. S. (1997). Maternal responsiveness and infant mental abilities: Specific predictive relations. *Infant Behavior & Development*, *20*(3), 283–296.

https://doi.org/10.1016/S0163-6383(97)90001-1

Bowlby, J. (1969). *Attachment and Loss: Vol. 1 Attachment* (Vol. 1). Basic Books.

Bretherton, I. (2013). Revisiting Mary Ainsworth's conceptualization and assessments of maternal sensitivity-insensitivity. *Attachment & Human Development*, *15*(5–6), 460–484. https://doi.org/10.1080/14616734.2013.835128

Briggs-Gowan, M. J., & Carter, A. S. (2006). BITSEA Brief Infant-Toddler Social and Emotional Assessment Examiner's Manual. San Antonio, TX: Harcourt Assessment, Inc.

Bruce, J., Gunnar, M. R., Pears, K. C., & Fisher, P. A. (2013). Early adverse care, stress neurobiology, and prevention science: Lessons learned. *Prevention Science*, *14*, 247–256. https://doi.org/10.1007/s11121-012-0354-6

Cabrera, N. J. (2022). Commentary: Recognizing our similarities and celebrating our differences – parenting across cultures as a lens toward social justice and equity. *Journal of Child Psychology and Psychiatry*, *63*(4), 480–483. https://doi.org/10.1111/jcpp.13600

Calzada, E. J., Huang, K.-Y., Anicama, C., Fernandez, Y., & Brotman, L. M. (2012). Test of a cultural framework of parenting with Latino families of young children. *Cultural Diversity & Ethnic Minority Psychology*, *18*(3), 285–296. https://doi.org/10.1037/a0028694

Casillas, M. & Scaff, C. (2021). Analyzing contingent interactions in R with chattr. In T. Fitch, C. Lamm, H. Leder, & K. Tebmar-Raible (Eds.), *Proceedings of the 43rd Annual Virtual Meeting of the Cognitive Science Society* (CogSci 2021) (pp. 2540 – 2546). Vienna, AT: Cognitive Science Society.

Chamberlain, P., & Reid, J. B. (1987). Parent observation and report of child symptoms. *Behavioral Assessment*, *9*(1), 97–109.

Chelune, G. J., & Baer, R. A. (1986). Developmental norms for the Wisconsin Card Sorting test. *Journal of Clinical and Experimental Neuropsychology*, *8*(3), 219–228. https://doi.org/10.1080/01688638608401314

Cicchetti, D., Rogosch, F. A., & Toth, S. L. (2006). Fostering secure attachment in infants in maltreating families through preventive interventions. *Development and Psychopathology*, *18*(3), 623–649. https://doi.org/10.1017/s0954579406060329

Coleman, P., & Karraker, K. (2003). Maternal self-efficacy beliefs, competence in parenting, and toddlers' behavior and developmental status. *Infant Mental Health Journal*, *24*, 126–148. https://doi.org/10.1002/imhj.10048

Cooper, G., Hoffman, K., Powell, B., & Marvin, R. (2005). The Circle of Security Intervention: differential diagnosis and differential treatment. In *Enhancing early attachments: Theory, research, intervention, and policy* (pp. 127–151). Guilford Press.

Cristia, A., Lavechin, M., Scaff, C., Soderstrom, M., Rowland, C., Rasanen, O., Bunce, J., & Begelson, E. (2021). A thorough evaluation of the Language Environment Analysis (LENA) system. *Behavioral Research Methods, 53*(2), 467-486. https://doi.org/10.3758/s13428-020-01393-5

Davidov, M., Bar-Tuvia, S. M., Polacheck-Benishti, N., & Grusec, J. E. (2022). Two forms of mother–child reciprocity and their links to children's cooperativeness. *Social Development*, *31*(2), 339–355. https://doi.org/10.1111/sode.12558

Davidov, M., & Grusec, J. E. (2006). Untangling the links of parental responsiveness to distress and warmth to child outcomes. *Child Development*, *77*, 44–58. https://doi.org/10.1111/j.1467-8624.2006.00855.x

Dettori, J. R., Norvell, D. C., & Chapman, J. R. (2018). The sin of missing data: Is all forgiven

by way of imputation? *Global Spine Journal*, *8*(8), 892–894. https://doi.org/10.1177/2192568218811922

Donnelly, S., & Kidd, E. (2021). The longitudinal relationship between conversational turn-taking and vocabulary growth in early language development. *Child Development*, *92*(2), 609–625. https://doi.org/10.1111/cdev.13511

Dozier, M. (2003). Attachment-based treatment for vulnerable children. *Attachment & Human Development*, *5*(3), 253–257. https://doi.org/10.1080/1461673031000159615l

Dozier, M., Albus, K., Fisher, P. A., & Sepulveda, S. (2002). Interventions for foster parents: Implications for developmental theory. *Development and Psychopathology*, *14*(4), 843–860. https://doi.org/10.1017/S0954579402004091

Dozier, M., & Bernard, K. (2017). Attachment and Biobehavioral Catch-up: Addressing the needs of infants and toddlers exposed to inadequate or problematic caregiving. *Current Opinion in Psychology*, *15*, 111–117. https://doi.org/10.1016/j.copsyc.2017.03.003

Dozier, M., Peloso, E., Lewis, E., Laurenceau, J.-P., & Levine, S. (2008). Effects of an attachment-based intervention on the cortisol production of infants and toddlers in foster care. *Development and Psychopathology*, *20*(3), 845–859. https://doi.org/10.1017/S0954579408000400

Dunst, C. J., & Kassow, D. Z. (2008). Caregiver sensitivity, contingent social responsiveness, and secure infant attachment. *Journal of Early and Intensive Behavior Intervention*, *5*(1), 40–56. https://doi.org/10.1037/h0100409

Dunst, C., Lowe, L. W., & Bartholomew, P. (1989). Contingent social responsiveness, family ecology, and infant communicative competence. *NSSLHA Journal*, *17*, 39–49. https://doi.org/10.1044/nsshla_17_39

Dyer, N., Owen, M. T., & Caughy, M. O. (2014). Ethnic differences in profiles of mother–child interactions and relations to emerging school readiness in African American and Latin American children. *Parenting, Science and Practice*, *14*(3–4), 175–194. https://doi.org/10.1080/15295192.2014.972756

Eyberg, S., & Bussing, R. (2010). Parent–Child Interaction Therapy for preschool children with conduct problems. In *Clinical Handbook of Assessing and Treating Conduct Problems in Youth* (pp. 139–162). https://doi.org/10.1007/978-1-4419-6297-3_6

Eyberg, S. M., Funderburk, B. W., Hembree-Kigin, T. L., McNeil, C. B., Querido, J. G., & Hood, K. K. (2001). Parent-child interaction therapy with behavior problem children: One and two year maintenance of treatment effects in the family. *Child & Family Behavior Therapy*, *23*(4), 1–20. https://doi.org/10.1300/J019v23n04_01

Farrell, A. K., Simpson, J. A., Carlson, E. A., Englund, M. M., & Sung, S. (2017). The impact of stress at different life stages on physical health and the buffering effects of maternal sensitivity. *Health Psychology*, *36*(1), 35–44. https://doi.org/10.1037/hea0000424

Feldstein, S. N., Keller, F. R., Portman, R. E., Durham, R. L., Klebe, K. J., & Davis, H. P. (1999). A comparison of computerized and standard versions of the Wisconsin Card Sorting Test. *Clinical Neuropsychologist*, *13*(3), 303–313. https://doi.org/10.1076/clin.13.3.303.1744

Ferjan Ramírez, N., Hippe, D. S., & Kuhl, P. K. (2021). Comparing automatic and manual measures of parent-infant conversational turns: A word of caution. *Child Development*, *92*(2), 672–681. https://doi.org/10.1111/cdev.13495

Ferjan Ramírez, N., Lytle, S. R., & Kuhl, P. K. (2020). Parent coaching increases conversational turns and advances infant language development. *Proceedings of the National Academy*

of Sciences of the United States of America, *117*(7), 3484–3491. https://doi.org/10.1073/pnas.1921653117

Fisher, P. A., Frenkel, T. I., Noll, L. K., Berry, M., & Yockelson, M. (2016). Promoting healthy child development via a two-generation translational neuroscience framework: The Filming Interactions to Nurture Development video coaching program. *Child Development Perspectives*, *10*(4), 251–256. https://doi.org/10.1111/cdep.12195

Fisher, P. A., & Gilliam, K. S. (2012). Research into theory into practice: An overview of family based interventions for child antisocial behavior developed at the Oregon Social Learning Center. *Clinica y Salud*, *23*(3), 247–259. https://doi.org/10.5093/cl2012a16

Fisher, P. A., & Kim, H. K. (2007). Intervention effects on foster preschoolers' attachment-related behaviors from a randomized trial. *Prevention Science: The Official Journal of the Society for Prevention Research*, *8*(2), 161–170. https://doi.org/10.1007/s11121-007-0066-5

Fisher, P. A., & Skowron, E. A. (2017). Social-learning parenting intervention research in the era of translational neuroscience. *Current Opinion in Psychology*, *15*, 168–173. https://doi.org/10.1016/j.copsyc.2017.02.017

Fisher, P. A., Stoolmiller, M., Gunnar, M. R., & Burraston, B. O. (2007). Effects of a therapeutic intervention for foster preschoolers on diurnal cortisol activity. *Psychoneuroendocrinology*, *32*(8–10), 892–905. https://doi.org/10.1016/j.psyneuen.2007.06.008

Flannery, J. E., Beauchamp, K. G., & Fisher, P. A. (2017). The role of social buffering on chronic disruptions in quality of care: Evidence from caregiver-based interventions in foster children. *Social Neuroscience*, *12*(1), 86–91.

https://doi.org/10.1080/17470919.2016.1170725

Fok, C. C. T., & Henry, D. (2015). Increasing the sensitivity of measures to change. *Prevention Science*, *16*(7), 978–986. https://doi.org/10.1007/s11121-015-0545-z

Forgatch, M. S., & Kjøbli, J. (2016). Parent Management Training-Oregon Model: Adapting intervention with rigorous research. *Family Process*, *55*(3), 500–513. https://doi.org/10.1111/famp.12224

Fukkink, R. G. (2008). Video feedback in widescreen: A meta-analysis of family programs. *Clinical Psychology Review*, *28*(6), 904–916. https://doi.org/10.1016/j.cpr.2008.01.003

Funderburk, B. W., & Eyberg, S. (2011). Parent–child interaction therapy. In *History of psychotherapy: Continuity and change, 2nd ed* (pp. 415–420). American Psychological Association. https://doi.org/10.1037/12353-021

Gardner, F. (2000). Methodological issues in the direct observation of parent–child interaction: Do observational findings reflect the natural behavior of participants? *Clinical Child and Family Psychology Review*, *3*(3).

Gardner, F., & Leijten, P. (2017). Incredible Years parenting interventions: Current effectiveness research and future directions. *Current Opinion in Psychology*, *15*, 99–104. https://doi.org/10.1016/j.copsyc.2017.02.023

Gilkerson, J., Richards, J. A., Warren, S. F., Montgomery, J. K., Greenwood, C. R., Kimbrough, O. D., Hansen, J. H. L., & Paul, T. D. (2017). Mapping the early language environment using all-day recordings and automated analysis. *American Journal of Speech-Language Pathology*, *26*(2), 248–265. https://doi.org/10.1044/2016_AJSLP-15-0169

Giuliani, N. R., Beauchamp, K. G., Noll, L. K., & Fisher, P. A. (2019). A preliminary study investigating maternal neurocognitive mechanisms underlying a child-supportive

parenting intervention. *Frontiers in Behavioral Neuroscience*, *13*. https://doi.org/10.3389/fnbeh.2019.00016

Golinkoff, R. M., Can, D. D., Soderstrom, M., & Hirsh-Pasek, K. (2015). (Baby)Talk to me: The social context of infant-directed speech and its effects on early language acquisition. *Current Directions in Psychological Science*, *24*(5), 339–344. https://doi.org/10.1177/0963721415595345

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigi-type card-sorting problem. *Journal of Experimental Psychology, 38*(4), 404-411.

Guild, D. J., Toth, S. L., Handley, E. D., Rogosch, F. A., & Cicchetti, D. (2017). Attachment security mediates the longitudinal association between child-parent psychotherapy and peer relations for toddlers of depressed mothers. *Development and Psychopathology*, *29*(2), 587–600. https://doi.org/10.1017/S0954579417000207

Hayes, A. F. (2012). PROCESS: A versatile computational tool for observed variable mediation, moderation and conditional process modeling. [White paper]. Retrieved from http://www.afhayes.com/public/process2012.pdf.

Heaton, R. K. (1993). Wisconsin card sorting test computer version 2.0. Odessa, FL: Psychological Assessment Resources.

Helander, M., Asperholm, M., Wetterborg, D., Öst, L.-G., Hellner, C., Herlitz, A., & Enebrink, P. (2024). The efficacy of Parent Management Training with or without involving the child in the treatment among children with clinical levels of disruptive behavior: A meta-analysis. *Child Psychiatry & Human Development*, *55*(1), 164–181. https://doi.org/10.1007/s10578-022-01367-y

Hirsh-Pasek, K., & Burchinal, M. (2006). Mother and caregiver sensitivity over time: Predicting language and academic outcomes with variable- and person-centered approaches. *Merrill-Palmer Quarterly*, *52*(3), 449–485. https://doi.org/10.1353/mpq.2006.0027

Hutchings, J., Williams, M. E., & Leijten, P. (2023). Attachment, behavior problems and interventions. *Frontiers in Child and Adolescent Psychiatry*, *2*. https://doi.org/10.3389/frcha.2023.1156407

Imhof, A. (2019). *How do we measure responsive caregiving? An evaluation of video-coding instruments to quantify parent-child interaction.* [Unpublished Master's Thesis].

Imhof, A., Liu, S., Schlueter, L., Phu, T., Watamura, S., & Fisher, P. (2023). Improving children's expressive language and auditory comprehension through responsive caregiving: Evidence from a randomized controlled trial of a strength-based video-coaching intervention. *Prevention Science*, *24*(1), 84–93.

Imhof, A. Peake, S., & Fisher, P. (2023, March). *FLO: A Novel video-coding tool to evaluate responsive caregiver-child interactions.* Poster presented at the 2023 Society of Research on Child Development Biennial Meeting. Salt Lake City, UT.

Imhof, A., Watamura, S., Fisher, P., Moreno, A., Hurwhich-Reiss, E., Schleuter, L., … Brown, S. M. (2019, August 13). Denver ACF Pre-Registration. Retrieved from osf.io/fdgv5

Johnston, C., & Mash, E. J. (1989). A measure of parenting satisfaction and efficacy. *Journal of Clinical Child Psychology*, *18*(2), 167–175. https://doi.org/10.1207/s15374424jccp1802_8

Juffer, F., Bakermans-Kranenburg, M. J., & van IJzendoorn, M. H. (2017). Pairing attachment theory and social learning theory in video-feedback intervention to promote positive parenting. *Current Opinion in Psychology*, *15*, 189–194.

https://doi.org/10.1016/j.copsyc.2017.03.012

Juffer, F., Bakermans-Kranenburg, M. J., & Van Ijzendoorn, M. H. (2018). Video-feedback intervention to promote positive parenting and sensitive discipline: Development and meta-analytic evidence for its effectiveness. In *Handbook of attachment-based interventions* (pp. 1–26). The Guilford Press.

Kazdin, A. E. (2005). *Parent management training: Treatment for oppositional, aggressive, and antisocial behavior in children and adolescents* (pp. ix, 410). Oxford University Press.

Keller, H. (2018). Universality claim of attachment theory: Children's socioemotional development across cultures. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(45), 11414–11419. https://doi.org/10.1073/pnas.1720325115

Kim, J. H. (2019). Multicollinearity and misleading statistical results. *Korean Journal of Anesthesiology*, *72*(6), 558–569. https://doi.org/10.4097/kja.19087

Kim-Cohen, J., Moffitt, T. E., Caspi, A., & Taylor, A. (2004). Genetic and environmental processes in young children's resilience and vulnerability to socioeconomic deprivation. *Child Development*, *75*(3), 651–668. https://doi.org/10.1111/j.1467-8624.2004.00699.x

King, G., Nielsen, R., Coberley, C., Pope, J. E., & Wells, A. (2011). Avoiding randomization failure in program evaluation, with application to the Medicare Health Support program. *Population Health Management*, *14 Suppl 1*, S11-22. https://doi.org/10.1089/pop.2010.0074

Kochanska, G., Forman, D. R., & Coy, K. C. (1999). Implications of the mother-child relationship in infancy socialization in the second year of life. *Infant Behavior & Development*, *22*(2), 249–265. https://doi.org/10.1016/S0163-6383(99)00009-0

Kochanska, G., Kim, S., Boldt, L. J., & Nordling, J. K. (2013). Promoting toddlers' positive social-emotional outcomes in low-income families: A play-based experimental study. *Journal of Clinical Child and Adolescent Psychology, 42*(5), 700–712. https://doi.org/10.1080/15374416.2013.782815

Kohlhoff, J., Lieneman, C., Cibralic, S., Traynor, N., & McNeil, C. B. (2022). Attachment-based parenting interventions and evidence of changes in toddler attachment patterns: An overview. *Clinical Child and Family Psychology Review*, *25*(4), 737–753. https://doi.org/10.1007/s10567-022-00405-4

Kuchirko, Y., Tafuro, L., & Tamis LeMonda, C. S. (2018). Becoming a communicative partner: infant contingent responsiveness to maternal language and gestures. *Infancy*, *23*(4), 558–576. https://doi.org/10.1111/infa.12222

Landry, S. H., Smith, K. E., & Swank, P. R. (2006). Responsive parenting: Establishing early foundations for social, communication, and independent problem-solving skills. *Developmental Psychology*, *42*(4), 627–642. https://doi.org/10.1037/0012-1649.42.4.627

Landry, S. H., Smith, K. E., Swank, P. R., Assel, M. A., & Vellet, S. (2001). Does early responsive parenting have a special importance for children's development or is consistency across early childhood necessary? *Developmental Psychology*, *37*(3), 387–403. https://doi.org/10.1037//0012-1649.37.3.387

Lansford, J. E. (2022). Annual research review: Cross-cultural similarities and differences in parenting. *Journal of Child Psychology and Psychiatry*, *63*(4), 466–479. https://doi.org/10.1111/jcpp.13539

Lewinsohn, P. M., Seeley, J. R., Roberts, R. E., & Allen, N. B. (1997). Center for Epidemiologic Studies Depression Scale (CES-D) as a screening instrument for depression among

community-residing older adults. *Psychology and Aging*, *12*(2), 277–287. https://doi.org/10.1037//0882-7974.12.2.277

Lieberman, A. F., Diaz, M., Castro, G., & Oliver Bucio, G. (2020). *Make room for baby: Perinatal Child-Parent Psychotherapy to repair trauma and promote attachment.* Guilford Press.

Lieberman, A. F., Ippen, C. G., & Horn, P. V. (2006). Child-Parent Psychotherapy: 6-Month follow-up of a randomized controlled trial. *Journal of the American Academy of Child & Adolescent Psychiatry*, *45*(8), 913–918. https://doi.org/10.1097/01.chi.0000222784.03735.92

Lieberman, A. F. & Van Horn, P. (2004). Assessment and treatment of young children exposed to traumatic events. In J. D. Osofsky (Ed.), *Young children and trauma: Intervention and treatment.* (pp. 111-138). Guilford Press.

Lieneman, C. C., Brabson, L. A., Highlander, A., Wallace, N. M., & McNeil, C. B. (2017). Parent-Child Interaction Therapy: Current perspectives. *Psychology Research and Behavior Management*, *10*, 239–256. https://doi.org/10.2147/PRBM.S91200

Lind, T., Bernard, K., Ross, E., & Dozier, M. (2014). Intervention effects on negative affect of CPS-referred children: Results of a randomized clinical trial. *Child Abuse & Neglect*, *38*(9), 1459–1467. https://doi.org/10.1016/j.chiabu.2014.04.004

Little, R. J. A. (1988). A test of missing completely at random for multivariate data with missing values. *Journal of the American Statistical Association*, *83*(404), 1198–1202. https://doi.org/10.1080/01621459.1988.10478722

Little, T., Jorgensen, T., Lang, K., & Moore, E. W. (2014). On the joys of missing data. *Journal of Pediatric Psychology, 39*(2), 151-162. https://doi.org/10.1093/jpepsy/jst048

Liu, S., Fisher, P. A., Schlueter, L. J., Phu, T., Gunnar, M. R., & Watamura, S. E. (2021). A brief video-coaching intervention buffers young children's vulnerability to the impact of caregivers' depressive symptoms: Examination of differential susceptibility. *Development and Psychopathology*, *33*(5), 1685–1700.

Liu, S., Phu, T., Dominguez, A., Hurwich-Reiss, E., McGee, D., Watamura, S., & Fisher, P. (2021). Improving caregiver self-efficacy and children's behavioral outcomes via a brief strength-based video coaching intervention: Results from a randomized controlled trial. *Prevention Science*, 1–14.

Lowe, J., Qeadan, F., Leeman, L., Shrestha, S., Stephen, J. M., & Bakhireva, L. N. (2017). The effect of prenatal substance use and maternal contingent responsiveness on infant affect. *Early Human Development*, *115*, 51–59. https://doi.org/10.1016/j.earlhumdev.2017.09.013

Lyons-Ruth, K. (1996). Attachment relationships among children with aggressive behavior problems: The role of disorganized early attachment patterns. *Journal of Consulting and Clinical Psychology*, *64*(1), 64–73. https://doi.org/10.1037/0022-006X.64.1.64

Maccoby, E. E. (1999). The uniqueness of the parent–child relationship. In *Relationships as developmental contexts* (pp. 157–175). Lawrence Erlbaum Associates Publishers.

Maccoby, E. E., & Martin, J. A. (1983), Socialization in the context of the family: Parent-child interaction. In E. M. Hetherington, & P. H. Mussen (Eds.), *Handbook of Child Psychology, Vol 4. Socialization, Personality, and Social Development* (pp. 1 – 101). Wiley.

Maciel, L., Gomis-Pomares, A., Day, C., & Basto-Pereira, M. (2023). Cross-cultural adaptability of parenting interventions designed for childhood behavior problems: A meta-analysis.

Clinical Psychology Review, *102*, 102274. https://doi.org/10.1016/j.cpr.2023.102274

McDonough, S. (1995). Promoting positive early parent-infant relationships through Interaction Guidance. *Infant Psychiatry*, *4*(3), 661–672.

McDorman, S. A., Imhof, A. Ramirez, A. G., Taylor, E. K., Abdurokhmonova, G., Roche, E. C., & Romeo, R. R. (2024, June). *Contextual moderators of caregiver-child conversational turns across lab-based and naturalistic settings.* Poster to be presented at the International Society for the Study of Behavioral Development 27th Biennial Meeting. Lisbon, Portugal.

McElwain, N. L., & Booth-Laforce, C. (2006). Maternal sensitivity to infant distress and nondistress as predictors of infant-mother attachment security. *Journal of Family Psychology: JFP: Journal of the Division of Family Psychology of the American Psychological Association (Division 43)*, *20*(2), 247–255. https://doi.org/10.1037/0893-3200.20.2.247

Mendez, M., Durtschi, J., Neppl, T. K., & Stith, S. M. (2016). Corporal punishment and externalizing behaviors in toddlers: The moderating role of positive and harsh parenting. *Journal of Family Psychology*, *30*(8), 887–895. https://doi.org/10.1037/fam0000187

Mesman, J., Minter, T., Angnged, A., Cissé, I. A. H., Salali, G. D., & Migliano, A. B. (2018). Universality without uniformity: A culturally inclusive approach to sensitive responsiveness in infant caregiving. *Child Development*, *89*(3), 837–850. https://doi.org/10.1111/cdev.12795

Michelson, D., Davenport, C., Dretzke, J., Barlow, J., & Day, C. (2013). Do evidence-based interventions work when tested in the "real world?" A systematic review and meta-analysis of parent management training for the treatment of child disruptive behavior.

Clinical Child and Family Psychology Review, *16*(1), 18–34. https://doi.org/10.1007/s10567-013-0128-0

Miller, G. E., Lachman, M. E., Chen, E., Gruenewald, T. L., Karlamangla, A. S., & Seeman, T. E. (2011). Pathways to resilience: Maternal nurturance as a buffer against the effects of childhood poverty on metabolic syndrome at midlife. *Psychological Science*, *22*(12), 1591–1599. https://doi.org/10.1177/0956797611419170

Mouton, B., Loop, L., Roskam, I., & Stievenart, M. (2018). Parenting programs to reduce young children's externalizing behavior: A meta-analytic review of their behavioral or cognitive orientation. *Child & Family Behavior Therapy*, *online publication*. https://doi.org/10.1080/07317107.2018.1477348

Mundy, P., Delgado, C., Goldstein, J., Parlade, M., Hogan, A., Seibert, J., & Mundy, D. (2003). *Early social communication scales (ESCS)*. *305*.

National Scientific Council on the Developing Child. (2004). *Young children develop in an environment of relationships: Working paper No. 1*. Retrieved from www.developingchild.harvard.edu.

National Scientific Council on the Developing Child. (2005/2014). *Excessive stress disrupts the architecture of the developing brain: Working paper No. 3.* Retrieved from www.developingchild.harvard.edu.

Narayan, A. J., Lieberman, A. F., & Masten, A. S. (2021). Intergenerational transmission and prevention of adverse childhood experiences (ACEs). *Clinical Psychology Review*, *85*, 101997. https://doi.org/10.1016/j.cpr.2021.101997

O'Connor, T. G., Matias, C., Futh, A., Tantam, G., & Scott, S. (2013). Social learning theory parenting intervention promotes attachment-based caregiving in young children:

Randomized clinical trial. *Journal of Clinical Child and Adolescent Psychology: The Official Journal for the Society of Clinical Child and Adolescent Psychology, American Psychological Association, Division 53*, *42*(3), 358–370. https://doi.org/10.1080/15374416.2012.723262

Owen, A. M., McMillan, K. M., Laird, A. R., & Bullmore, E. (2005). N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Human Brain Mapping*, *25*(1), 46–59. https://doi.org/10.1002/hbm.20131

Parpal, M., & Maccoby, E. E. (1985). Maternal responsiveness and subsequent child compliance. *Child Development*, *56*(5), 1326–1334. https://doi.org/10.2307/1130247

Patterson, G. R. (1982). *Coercive Family Process* (Vol. 3). Castalia Publishing Company.

Patterson, G. R. & Fisher, P. A. (2002). Recent developments in our understanding of parenting: Bidirectional effects, causal models, and the search for parsimony. In M. H. Bornstein (Ed.), *Handbook of parenting: Practical and applied parenting.* (2nd Ed., pp – 59-88). Lawrence Erlbaum Associates Publishers.

Patterson, G. R., Forgatch, M. S., & DeGarmo, D. S. (2010). Cascading effects following intervention. *Development and Psychopathology*, *22*(4), 949–970. https://doi.org/10.1017/S0954579410000568

Patterson, G. R., & Reid, J. B. (1984). Social interactional processes within the family: The study of the moment-by-moment family transactions in which human social development is imbedded. *Journal of Applied Developmental Psychology*, *5*(3), 237–262. https://doi.org/10.1016/0193-3973(84)90021-2

Peña, E. D., Bedore, L. M., & Kester, E. S. (2016). Assessment of language impairment in bilingual children using semantic tasks: Two languages classify better than one.

International Journal of Language & Communication Disorders, *51*(2), 192–202. https://doi.org/10.1111/1460-6984.12199

Peredo, T. N., Dillehay, K. M., & Kaiser, A. P. (2020). Latino caregivers' interactions with their children with language delays: A comparison study. *Topics in Early Childhood Special Education*, *40*(1), 52–63. https://doi.org/10.1177/0271121419900269

Pidano, A. E., & Allen, A. R. (2015). The Incredible Years series: A review of the independent research base. *Journal of Child and Family Studies*, *24*(7), 1898–1916. https://doi.org/10.1007/s10826-014-9991-7

Raby, K. L., Roisman, G. I., Fraley, R. C., & Simpson, J. A. (2015). The enduring predictive significance of early maternal sensitivity: Social and academic competence through age 32 Years. *Child Development*, *86*(3), 695–708. https://doi.org/10.1111/cdev.12325

Radloff, L. S. (1977). The CES-D Scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement*, *1*(3), 385–401. https://doi.org/10.1177/014662167700100306

Raver, C. C. (1996). Relations between social contingency in mother-child interaction and 2-year-olds' social competence. *Developmental Psychology*, *32*(5), 850–859. https://doi.org/10.1037/0012-1649.32.5.850

Ríos, M., Zekri, S., Alonso-Esteban, Y., & Navarro-Pardo, E. (2022). Parental stress assessment with the Parenting Stress Index (PSI): A systematic review of its psychometric properties. *Children*, *9*(11), 1649. https://doi.org/10.3390/children9111649

Romeo, R. R., Leonard, J. A., Grotzinger, H. M., Robinson, S. T., Takada, M. E., Mackey, A. P., Scherer, E., Rowe, M. L., West, M. R., & Gabrieli, J. D. E. (2021). Neuroplasticity associated with changes in conversational turn-taking following a family-based

intervention. *Developmental Cognitive Neuroscience*, *49*, 100967. https://doi.org/10.1016/j.dcn.2021.100967

Romeo, R. R., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Rowe, M. L., & Gabrieli, J. D. E. (2018). Beyond the 30-million-word gap: Children's conversational exposure is associated with language-related brain function. *Psychological Science*, *29*(5), 700–710. https://doi.org/10.1177/0956797617742725

Romeo, R. R., Segaran, J., Leonard, J. A., Robinson, S. T., West, M. R., Mackey, A. P., Yendiki, A., Rowe, M. L., & Gabrieli, J. D. E. (2018). Language exposure relates to structural neural connectivity in childhood. *Journal of Neuroscience*, *38*(36), 7870–7877. https://doi.org/10.1523/JNEUROSCI.0484-18.2018

Roth, D. L., Ackerman, M. L., Okonkwo, O. C., & Burgio, L. D. (2008). The four-factor model of depressive symptoms in dementia caregivers: A structural equation model of ethnic differences. *Psychology and Aging*, *23*(3), 567–576. https://doi.org/10.1037/a0013287

Samdan, G., Kiel, N., Petermann, F., Rothenfußer, S., Zierul, C., & Reinelt, T. (2020). The relationship between parental behavior and infant regulation: A systematic review. *Developmental Review*, *57*, 100923. https://doi.org/10.1016/j.dr.2020.100923

Schindler, H. S., Fisher, P. A., Ola, C., & Campbell, C. J. (2021). Filmando Interacciones Para Nutrir El Desarrollo: A randomized pilot of a strength-based video coaching program with Mexican American fathers. *Early Education and Development, 33*(8), 1420-1438. https://doi.org/10.1080/10409289.2021.1963648

Schindler, H. S., Fisher, P. A., & Shonkoff, J. P. (2017). From innovation to impact at scale: lessons learned from a cluster of research-community partnerships. *Child Development*, *88*(5), 1435–1446. https://doi.org/10.1111/cdev.12904

Shin, H., Park, Y.-J., Ryu, H., & Seomun, G.-A. (2008). Maternal sensitivity: A concept analysis. *Journal of Advanced Nursing*, *64*(3), 304–314. https://doi.org/10.1111/j.1365-2648.2008.04814.x

Shonkoff, J. P., & Bales, S. N. (2011). Science does not speak for itself: Translating child development research for the public and its policymakers. *Child Development*, *82*(1), 17–32. https://doi.org/10.1111/j.1467-8624.2010.01538.x

Shonkoff, J. P., Phillips, D. A. (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. National Academies Press.

Skowron, E. A., Nekkanti, A. K., Skoranski, A. M., Scholtes, C. M., Lyons, E. R., Mills, K. L., Bard, D., Rock, A., Berkman, E., Bard, E., & Funderburk, B. W. (2024). Randomized trial of Parent–Child Interaction Therapy improves child-welfare parents' behavior, self-regulation, and self-perceptions. *Journal of Consulting and Clinical Psychology*, *92*(2), 75–92. https://doi.org/10.1037/ccp0000859

Spitzer, R. L., Kroenke, K., Williams, J. B. W., & Löwe, B. (2006). A brief measure for assessing generalized anxiety disorder: The GAD-7. *Archives of Internal Medicine*, *166*(10), 1092–1097. https://doi.org/10.1001/archinte.166.10.1092

Sprang, G. (2009). The efficacy of a relational treatment for maltreated children and their families. *Child and Adolescent Mental Health*, *14*(2), 81–88. https://doi.org/10.1111/j.1475-3588.2008.00499.x

Stams, G.-J. J. M., Juffer, F., & van IJzendoorn, M. H. (2002). Maternal sensitivity, infant attachment, and temperament in early childhood predict adjustment in middle childhood: The case of adopted children and their biologically unrelated parents. *Developmental Psychology*, *38*(5), 806–821. https://doi.org/10.1037//0012-1649.38.5.806

Stronach, E. P., Toth, S. L., Rogosch, F., & Cicchetti, D. (2013). Preventive interventions and sustained attachment security in maltreated children. *Development and Psychopathology*, *25*(4 Pt 1), 919–930. https://doi.org/10.1017/S0954579413000278

Sutton, T. E. (2019). Review of attachment theory: Familial predictors, continuity and change, and intrapersonal and relational outcomes. *Marriage & Family Review*, *55*(1), 1–22. https://doi.org/10.1080/01494929.2018.1458001

Tamis-LeMonda, C. S., & Bornstein, M. H. (2002). Maternal responsiveness and early language acquisition. In R. V. Kail & H. W. Reese (Eds.), *Advances in Child Development and Behavior* (Vol. 29, pp. 89–127). JAI. https://doi.org/10.1016/S0065-2407(02)80052-0

Tamis-LeMonda, C. S., Bornstein, M. H., & Baumwell, L. (2001). Maternal responsiveness and children's achievement of language milestones. *Child Development*, *72*(3), 748–767.

Tamis-LeMonda, C. S., Caughy, M. O., Rojas, R., Bakeman, R., Adamson, L. B., Pacheco, D., Owen, M. T., Suma, K., & Pace, A. (2020). Culture, parenting, and language: Respeto in Latine mother–child interactions. *Social Development (Oxford, England)*, *29*(3), 689–712. https://doi.org/10.1111/sode.12430

Tamis-LeMonda, C. S., Kuchirko, Y., Luo, R., Escobar, K., & Bornstein, M. (2017). Power in methods: Language to infants in structured and naturalistic contexts. *Developmental Science, 20*(6), 1-22. https://doi.org/10.1111/desc.12456

Tamis-LeMonda, C. S., Kuchirko, Y., & Song, L. (2014). Why is infant language learning facilitated by parental responsiveness? *Current Directions in Psychological Science*, *23*(2), 121–126. https://doi.org/10.1177/0963721414522813

Taylor, T. K., & Biglan, A. (1998). Behavioral family interventions for improving child-rearing: A review of the literature for clinicians and policy makers. *Clinical Child and Family*

Psychology Review, *1*(1), 41–60. https://doi.org/10.1023/a:1021848315541

Thomas, R., Abell, B., Webb, H. J., Avdagic, E., & Zimmer-Gembeck, M. J. (2017). Parent-Child Interaction Therapy: A meta-analysis. *Pediatrics*, *140*(3), e20170352. https://doi.org/10.1542/peds.2017-0352

Thomas, R., & Zimmer-Gembeck, M. J. (2011). Accumulating evidence for Parent-Child Interaction Therapy in the prevention of child maltreatment. *Child Development*, *82*(1), 177–192.

Toth, S. L., & Gravener, J. (2012). Review: Bridging research and practice: relational interventions for maltreated children. *Child and Adolescent Mental Health*, *17*(3), 131–138. https://doi.org/10.1111/j.1475-3588.2011.00638.x

Valentino, K. (2017). Relational interventions for maltreated children. *Child Development*, *88*(2), 359–367. https://doi.org/10.1111/cdev.12735

Van Aken, C., Junger, M., Verhoeven, M., Van Aken, M. A. G., & Deković, M. (2007). The interactive effects of temperament and maternal parenting on toddlers' externalizing behaviours. *Infant and Child Development*, *16*(5), 553–572. https://doi.org/10.1002/icd.529

Van Dam, N. T., & Earleywine, M. (2011). Validation of the Center for Epidemiologic Studies Depression Scale—Revised (CESD-R): Pragmatic depression assessment in the general population. *Psychiatry Research*, *186*(1), 128–132. https://doi.org/10.1016/j.psychres.2010.08.018

van der Klis, A., Adriaans, F., & Kager, R. (2023). Infants' behaviours elicit different verbal, nonverbal, and multimodal responses from caregivers during early play. *Infant Behavior & Development*, *71*, 1–17. https://doi.org/10.1016/j.infbeh.2023.101828

van Rijen, E. H. M., Gasanova, N., Boonstra, A. M., & Huijding, J. (2014). Psychometric qualities of the Short Form of the Self-efficacy for Parenting Tasks Index-Toddler Scale. *Child Psychiatry and Human Development*, *45*(4), 443–455. https://doi.org/10.1007/s10578-013-0414-6

Van Zeijl, J., Mesman, J., Van IJzendoorn, M. H., Bakermans-Kranenburg, M. J., Juffer, F., Stolk, M. N., Koot, H. M., & Alink, L. R. A. (2006). Attachment-based intervention for enhancing sensitive discipline in mothers of 1- to 3-year-old children at risk for externalizing behavior problems: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, *74*(6), 994–1005. https://doi.org/10.1037/0022-006X.74.6.994

Vance, A. J., & Brandon, D. H. (2017). Delineating among parenting confidence, parenting self-efficacy and competence. *ANS. Advances in Nursing Science*, *40*(4), E18–E37. https://doi.org/10.1097/ANS.0000000000000179

Vernon-Feagans, L., & Bratsch-Hines, M. E. (2013). Caregiver–child verbal interactions in child care: A buffer against poor language outcomes when maternal language input is less. *Early Childhood Research Quarterly*, *28*, 858–873. https://doi.org/10.1016/j.ecresq.2013.08.002

Wasik, B. A., & Jacobi-Vessels, J. L. (2017). Word play: Scaffolding language development through child-directed play. *Early Childhood Education Journal*, *45*(6), 769–776. https://doi.org/10.1007/s10643-016-0827-5

Waters, S. F., Hagan, M. J., Rivera, L., & Lieberman, A. F. (2015). Improvements in the child-rearing attitudes of Latina mothers exposed to interpersonal trauma predict greater maternal sensitivity toward their 6-month-old infants. *Journal of Traumatic Stress*, *28*(5), 426–433. https://doi.org/10.1002/jts.22043

Webster-Stratton, C., & Reid, M. J. (2010). The Incredible Years parents, teachers, and children training series: A multifaceted treatment approach for young children with conduct disorders. In *Evidence-based psychotherapies for children and adolescents, 2nd ed* (pp. 194–210). The Guilford Press.

Wiggins, J. L., Mitchell, C., Hyde, L. W., & Monk, C. S. (2015). Identifying early pathways of risk and resilience: The codevelopment of internalizing and externalizing symptoms and the role of harsh parenting. *Development and Psychopathology*, *27*(4pt1), 1295–1312. https://doi.org/10.1017/S0954579414001412

Wyatt Kaminski, J., Valle, L. A., Filene, J. H., & Boyle, C. L. (2008). A Meta-analytic review of components associated with parent training program effectiveness. *Journal of Abnormal Child Psychology*, *36*(4), 567–589. https://doi.org/10.1007/s10802-007-9201-9

Zajac, L., Raby, K. L., & Dozier, M. (2020). Sustained effects on attachment security in middle childhood: Results from a randomized clinical trial of the Attachment and Biobehavioral Catch-up (ABC) intervention. *Journal of Child Psychology and Psychiatry*, *61*(4), 417–424. https://doi.org/10.1111/jcpp.13146

Zayas, L. H. Borrego, J. Hr., & Domenech Rodriguez, M. M. (2009). Parenting interventions and Latino families: Research findings cultural adaptations, and future directions. In F. A. Villarruel, G. Carlo, J. M. Grau, M. Azmitia, N. J. Cabrera, & T. J. Chahin (Eds.), *Handbook of U.S. Latino Psychology: Developmental and Community-Based Perspectives* pp. 291 – 307). Sage Publications, Inc.

Zimmerman, I. L., Steiner, V. G., & Pond, R. E. (2011). *Preschool Language Scale, Fifth Edition*. https://doi.org/10.1037/t15141-000

Zumbach, J., Oster, A., Rademacher, A., & Koglin, U. (2022). Reliability and validity of

behavior observation coding systems in child maltreatment risk evaluation: A systematic review. *Journal of Child and Family Studies*, *31*(2), 545–562. https://doi.org/10.1007/s10826-021-02160-6