AMERICAN PSYCHOLOGICAL ASSOCIATION

Developmental Psychology

ISSN: 0012-1649

2022, Vol. 58, No. 5, 835–847
https://doi.org/10.1037/dev0001332

Positive Impacts of a Strengths-Based Family Program on Latino Kindergarteners' Narrative Language Abilities

Diana Leyva[1], Anna Shapiro[2], Gloria Yeomans-Maldonado[3], Christina Weiland[4], and Kathryn Leech[5]

[1] Department of Psychology, University of Pittsburgh
[2] School of Education and Human Development, University of Virginia
[3] Children's Learning Institute, The University of Texas Health Science Center at Houston
[4] School of Education, University of Michigan
[5] School of Education, University of North Carolina at Chapel Hill

Narrative language abilities are foundational to literacy development and are a culturally grounded measure of early literacy for Latino children. This study evaluates the impacts on narrative language abilities and the costs of a 4-week, strengths-based program that leverages two valued sociocultural practices with built-in benefits, personal narratives, and family food routines (e.g., grocery shopping), for improving Latino kindergarteners' learning outcomes in the United States. Two-hundred and 34 children (M age = 67 months; 51% girls; 13 schools) and their parents participated in a cluster randomized trial. Children produced personal narratives at three time points: pretest, end-of-treatment, and 5-month follow-up. Four narrative features were measured: narrative coherence, elaborations, word types, and literate language features. Large positive impacts were observed on all four narrative features at the end-of-treatment posttest (d = 1.21–1.76). There was suggestive evidence of moderate impacts on one narrative feature (i.e., narrative coherence) at the 5-month follow-up (d = .59). The costs required to implement the family program were relatively low. Findings highlight the potential value of implementing this strengths-based program in schools serving Latino kindergarteners using a rigorous evaluation of its effectiveness.

Keywords: Latino, intervention, strengths-based, narrative, language

Supplemental materials: https://doi.org/10.1037/dev0001332.supp

There is wide variability in the impacts of family interventions aimed at improving racially minoritized and linguistically diverse children's language and literacy outcomes (Dowdall et al., 2020; Manz et al., 2010). One factor that might explain this variability is the lack of alignment between the values and practices promoted by some interventions and those of the community they intend to serve. Many interventions focus on what parents "seemingly" lack (deficit approach) rather than leveraging what families are already doing to help their children learn (i.e., strengths-based approach; Perez-Brena et al., 2018). Another factor might be the lack of culturally grounded assessments (Fiestas & Pena, 2004). Intervention impacts are typically measured using standardized assessments, which assume certain content knowledge that racially minoritized and linguistically diverse children might not necessarily possess. Standardized assessments often fail to measure children's *ecocultural assets*, the competencies that racially minoritized and linguistically diverse children excel in, which are foundational to the development of other competencies (e.g., academic skills), and are sustained by the sociocultural practices and values of their communities (Perez-Brena et al., 2018).

While the importance of adopting a strengths-based approach is well established, very few interventions are designed with this approach in mind, and even fewer have used culturally grounded measures to evaluate intervention impacts (Melzi et al., 2019). This study addressed this gap by utilizing a cluster randomized controlled trial (RCT) design to evaluate the effects of Food For Thought (henceforth, FFT), a strengths-based and culturally responsive intervention that leverages two valued sociocultural

This article was published Online First March 17, 2022.
Diana Leyva https://orcid.org/0000-0002-8411-0850
Anna Shapiro https://orcid.org/0000-0001-9312-1300
Kathryn Leech https://orcid.org/0000-0003-1461-3855
This project was supported by the Brady Education Foundation, Davidson College, and the Institute of Education Sciences under Grant R305B150012 and R305B170015. We thank all children and families who participated in this project; Catherine Snow and the snowcats members; participants in the University of Michigan Education Policy Initiative's Education Policy Initiative's Causal Inference in Education Research Seminar; and the EdPolicyWorks Research Seminar at the University of Virginia for their helpful comments. We also thank Angela Febles, Yarelin Rivera, Danielle Mayall, and the undergraduate student research assistants who made this work possible. The study's design, hypotheses, and analysis plan were preregistered; see https://aspredicted.org/blind.php?x=62m22i. Data and study materials and analysis code are available upon request from the authors.

Correspondence concerning this article should be addressed to Diana Leyva, Department of Psychology, University of Pittsburgh, Sennott Square, 3rd Floor, 210 South Bouquet Street, Pittsburgh, PA 15260, United States. Email: dml114@pitt.edu

practices with built-in benefits, personal narratives (stories about personal experiences), and family food routines (e.g., grocery shopping, cooking), for improving Latino[1] kindergarteners' narrative language abilities, and assess its impacts via a culturally grounded assessment (i.e., narrative task).

Our work is grounded in Garcia-Coll et al.'s (1996) integrative model of developmental competencies in minority children, whereby social position factors (e.g., race, ethnicity) influence how families experience proximal ecologies (e.g., racial discrimination in schools), which in turn influence their practices (e.g., socialization goals and parenting strategies). These family practices (e.g., teach children how to prepare and respond to racial discrimination) may protect minority children from risk and promote positive outcomes (Garcia-Coll et al., 1996; Perez-Brena et al., 2018). Following this model, FFT focuses on Latino children's daily experiences (i.e., personal narratives and family food routines), the racial and ethnic values that facilitate the development of children's abilities in these contexts (i.e., sense of efficacy and competence in familiar, sociocultural practices, which results in more cognitively complex and engaging parent–child interactions), and the structural factors that can impede this development (e.g., lack familiarity with the U.S. educational system). In this evaluation, we also include a cost analysis exploring the resources needed to implement this program in the future.

Improving Children's Narrative Language Abilities Through Parent–Child Narratives

Narrative language abilities (e.g., the ability to tell stories) require integration of a variety of skills (e.g., vocabulary, grammar, syntax, morphology), and this integration is essential to reading comprehension (Florit & Cain, 2011). Children who start kindergarten with strong oral language abilities (including narrative language abilities) are more likely to develop strong reading comprehension skills and succeed academically than children who struggle with oral language abilities early on (Dickinson & Tabors, 2001). Narrative language abilities also require critical cognitive abilities such as planning and organizing the discourse production and monitoring the message for coherence (McCabe & Rollins, 1994). Narratives are a form of decontextualized talk, that is, talk about events, people, and objects that are removed from the immediate context (not present in here and now) (Rowe, 2013). Decontextualized talk in preschool is a strong predictor of school achievement in adolescence (Uccelli et al., 2019).

RCTs of family interventions promoting parent–child personal narratives have shown that is possible to increase parental use of strategies that effectively engage children in these narratives (e.g., asking open-ended questions such as what and who, following the child's lead, and using backchannelling responses such as uh-huh and yeah), and such increases lead to improvements in children's narrative language abilities. RCTs targeting mixed-income White parents and their toddlers and preschoolers in New Zealand and Canada have found enhancements in children's narrative coherence (i.e., ability to provide temporal and location context while telling the story) and elaborations (i.e., number of unique units of information provided) at the 1-year follow-up (Peterson et al., 1999; n = 20; Reese & Newcombe, 2007; n = 100). An RCT targeting middle-income White families and their preschoolers (n = 36) in the U.S. and promoting narratives during mealtimes and other forms of decontextualized talk found improvements in children's use of decontextualized talk, operationalized as number of utterances containing past and future talk and explanations, at the end-of-treatment posttest (Leech et al., 2018). Three small RCTs specifically targeted Latino families and their preschoolers in the United States. Two of them promoted parent–child personal narratives and found enhancements in the number of elaborations in children's narratives at the end-of-treatment posttest (Reese et al., 2010; n = 33; 33% were Latino) and the number of details in children's narratives at the 4-month posttest (Thierry & Sparks, 2013; n = 66). The third one promoted book-sharing and found positive impacts on the number of different words (word types) in children's narratives at the end-of-treatment posttest (Hammer & Sawyer, 2016; n = 73).

Overall, previous experimental evidence has either focused on White families, targeted younger children, and/or relied on small sample sizes. As such, the extent to which a strengths-based program promoting parent–child personal narratives can have a positive impact on Latino children beyond preschool remains unclear. A rigorous evaluation of such programs using culturally grounded assessments on a large sample of Latino children can yield a more precise estimate, that can in turn inform practice and policy.

The FFT Program

FFT Development

Latino children are the fastest growing minority population in the U.S.; 23% of school-age children are of Latino heritage (U.S. Census Bureau, 2019). In interviews with stakeholders in the school district where FFT was implemented, we learned that schools lacked culturally sustaining supports for Latino students. Thus, the lead author of the present study developed FFT to meet the needs of the underserved and growing Latino student population in schools. FFT is a 4-week family program that takes place in schools and targets kindergarten. This timing may be particularly optimal for Latino parents; due to inequality factors, fewer Latino children (60%) attend preschool than their Black and White peers (70%; Ansari, 2017). Accordingly, many Latino parents may be especially motivated to be involved in their children's education at kindergarten entry because it is the first formal schooling experience for many. Thus, kindergarten offers a unique window of opportunity for interventions to partner with these parents. FFT combines provision of information with onsite opportunities to practice, coaching and videoclips, based on best practices in behavioral change of adult population (Michie et al., 2009).

FFT Leverages Personal Narratives and Family Food Routines

Personal narratives and family food routines are unique practices that are already established in the ecocultural context of Latino family and have two built-in benefits. The first is the *high frequency* of these practices, which is sustained by cultural values and beliefs. Latino communities have a strong oral tradition; thus, families regularly engage in these practices, as they serve moral, religious, and social purposes (Hammer & Sawyer, 2016). Personal narratives are the most natural and familiar for Latino children, as children as young as 2 years of age produce them (Eisenberg,

[1] We use the term Latino (rather than Latinx, Latine, Latin*, or Latin@), to honor the way families preferred to be described in the study.

2002). During the preschool years, Latino children make substantial improvements in both the quantity (length) and the quality of their narratives (macrolevel and microlevel features; Melzi et al., 2011). As Latino preschoolers participate in family conversations as both active listeners and storytellers, their language, social and cognitive skills develop, and they begin to master macrolevel (ability to organize a coherent story and include key discourse elements) and microlevel features of narratives (breadth of vocabulary and syntactic knowledge). Advancements in the complexity of Latino children's narratives continue as they enter school. One study found that Latino narratives involved, on average, 21 different words (word types) and 36 total words (tokens) in kindergarten and 25 different words and 43 total words in first grade (Uccelli & Paez, 2007). Another study found that Latino children were, on average, 41% successful in kindergarten, and 50% successful in first grade at telling a coherent story (Miller et al., 2006).

Similarly, Latinos have the highest rates of shared mealtimes relative to any other ethnicity in the U.S. (6–7 nights per week; Murphey et al., 2014) and have high rates of child involvement in kitchen chores (Eisenberg, 2002; Evans et al., 2011). Food routines serve as a powerful vehicle to transmit and preserve Latino culture, as parents socialize their children into enacting *familismo* (strong sense of loyalty to family) and developing their identities as Latinos during these practices (Evans et al., 2011).

The second built-in benefit is the *type of parent–child interactions* that these practices afford. During these practices, Latino parents engage in more cognitively complex interactions with their children than in other practices. For example, Latino parents are more likely to ask questions, invite children to verbally contribute, and follow the child's lead during personal narratives than during book-sharing (Melzi et al., 2011). Similarly, Latino parents ask more questions requiring active thinking, provide more explanations, follow the child's lead, and encourage the child's contributions more during food-related activities (e.g., baking cookies) than non-food-related activities (e.g., book-sharing), due to their familiarity with these practices and sense of self-efficacy (e.g., Eisenberg, 2002; Kermani & Janes, 1999). Hence, through their participation in personal narratives and family food routines, Latino children are likely acquiring literacy skills indirectly (Melzi et al., 2011).

FFT's Theory of Change

FFT's goal is to accelerate the use of strategies that Latino parents already utilize with their children during preexisting sociocultural practices that have built-in benefits. Rather than imposing new behaviors or practices that might not be culturally relevant or meaningful for families, FFT aims to propel Latino families' strengths to foster children's narrative language abilities. To engage children in personal narratives during food routines, FFT promotes parental use of the three strategies that are known to be effective: asking open-ended questions, following the child's lead, and using backchannelling responses to help the child elaborate (e.g., Peterson et al., 1999; Reese & Newcombe, 2007). FFT's theory of change is that parents in schools assigned to FFT would increase their use of these strategies in personal narratives during food routines, and as a result, their children would have enhanced narrative language abilities, relative to children in schools assigned to an active control condition. Note that FFT also promotes other practices (e.g., authentic reading and writing and math talk); however, we focus on parent–child narratives because they are directly related to narrative language abilities, the target outcome of the current study.

Why Use Narrative Assessments?

In recent years, calls to move beyond standardized measures to assess language and literacy growth in racially minoritized and linguistically diverse children in the U.S. have increased (Fiestas & Pena, 2004). Standardized assessments assume certain content knowledge that Latino children might not possess (Fiestas & Pena, 2004). Narrative assessments require children to tell or retell a fictional story prompted by wordless picture (e.g., Hammer & Sawyer, 2016) or regular books (e.g., Reese et al., 2010), or a personal story based on a reminiscing prompt (e.g., Peterson et al., 1999). Narratives are audio recorded, transcribed, and scored based on macrolevel (e.g., narrative coherence) and microlevel features (e.g., word types). For Latino children, narrative assessments might be an ideal culturally grounded measure to assess language and literacy, as they provide critical information about several oral language and other cognitive abilities that are foundational for reading comprehension, while leveraging children's ecocultural assets. Furthermore, Latino children's performance on narrative assessments in kindergarten to third grade uniquely predict their reading achievement (Miller et al., 2006). Thus, this RCT evaluation of FFT relied on narrative assessments.

Previous FFT Evidence

We piloted FFT in 2014 (N = 10 families, one school) and conducted a feasibility study of FFT in 2015 (N = 68, three schools), yielding promising results (Leyva & Skorb, 2017). Children whose parents attended more FFT sessions had larger vocabularies from pretest to posttest, relative to children whose parents attended fewer FFT sessions. We then launched a cluster-randomized trial in 2018 (N = 248, 13 schools). Previously, we reported FFT impacts on distal outcomes (i.e., child vocabulary, literacy, math, executive function, and approaches to learning; Leyva et al., 2021). We found no impacts of FFT on children's vocabulary or literacy skills using standardized measures (i.e., Woodcock-Muñoz Language Survey Revised; Woodcock et al., 2005) at the end-of-treatment and 5-month follow-up. We found moderate effects on a nonstandardized measure of children's expressive vocabulary (IDELA, Save the Children, 2017) at the end-of-treatment (d = .57) and 5-month follow-up (d = .59), and suggestive evidence of moderate effects on approaches to learning and executive function at the 5-month follow-up (d = .38-.95). Here we examine FFT's impacts on children's narrative language abilities (i.e., proximal outcome).

Cost Analysis

Cost analysis involves describing the "ingredients" or resources needed to run a program (e.g., personnel, facilities, materials), estimating the prices for those resources, and calculating cost estimates (Institute of Educational Sciences [IES], 2020). The purpose of cost analysis is to help state and local education providers and funders make decisions about whether to adopt a new program or continue an existing program designed to improve children's academic outcomes. Cost analysis can also help providers determine whether and

how to run a program with fidelity and whether the resources are used as planned. Although this kind of analysis might be new for some researchers, it has been a standard practice for economists for quite some time and is quickly becoming a requirement in many grant opportunities, including some from the Institute of Education Sciences (IES, 2020).

For example, a cost analysis study of the Reading Partners, a program that uses volunteers to provide one-on-one tutoring to struggling readers in underserved elementary schools across the U.S., found that the financial and other resources necessarily to implement the program in schools were low ($710 per student), which highlighted the potential value of the program, not only from the point of view of its effectiveness (positive impacts on literacy outcomes were detected), but also from the point of view of its feasibility to be implemented in real-world settings (Jacob et al., 2016). A cost analysis study of Sound Partner, a literacy program specifically targeting kindergarteners, yielded similarly low cost ($791 per student; Hollands et al., 2016).

Importantly, none of the RCTs of family intervention programs previously discussed included a cost analysis; hence, the extent to which these programs are feasible to implement and sustainable over time is uncertain. Thus, to offer a more complete accounting of FFT, we analyzed the economic costs of the resources required to implement this strengths-based and culturally responsive family program. Cost information on FFT allows stakeholders to determine the feasibility of this program and better compare it to other alternatives for improving kindergarteners' language and literacy skills.

The Present Study

The present study evaluated the FFT program's impacts on children's narrative language abilities and the program's cost. We addressed the following research questions:

1. Does FFT have a positive impact on children's narrative language abilities at the end-of-treatment and 5-month follow-up?

2. What resources are needed to implement FFT as described in this study?

We expected positive impacts of FFT on children's narrative language abilities at the end-of-treatment and hypothesized that such effects would persist at the 5-month follow-up. Families might choose to continue implementing targeted practices beyond the program implementation, given that they are embedded in their everyday practices and sustained by their cultural beliefs and values. We explored the cost of implementing FFT.

Method

Procedure

Research Design

This study received research ethics committee approval from Davidson College, protocol # 2018-030, study title: Evaluation of the Effectiveness of the Food For Thought Program. The study's design, hypotheses and analysis plan were preregistered (Leyva et al., 2018; see https://aspredicted.org/blind.php?x=62m22i). We used a cluster-randomized design to estimate the impact of the FFT program on children's narrative language abilities. Students were clustered in schools, which were randomly assigned to the FFT intervention condition or an active control condition. The final sample involved 13 schools with 261 kindergarten children across two cohorts (N = 129 Cohort 1 in 2018; N = 132 Cohort 2 in 2019).

Program Characteristics of the FFT and Active Control Conditions

The FFT program took place in each treatment school in the fall of the kindergarten year. FFT involved four 90-min group sessions (one per week) scheduled at convenient times for parents and school staff (typically, during school hours). A team of bilingual facilitators delivered the group sessions (15 in total; two Latina group leaders who had a master's degree and 12 bilingual research assistants, three of whom were Latinos). Facilitators received training (i.e., 3-hr training) and coaching (i.e., were observed and received feedback during implementation) by a master trainer. There was one group leader and one to two research assistants per session. Facilitators delivered the sessions in the parents' preferred language (i.e., Spanish only or Spanish/English) with most of the sessions (95%) delivered in Spanish. Each 90-min session involved a "learning" (60 min) and a "practice" (30 min) component. During "learning," parents watched and discussed video clips of other Latino parents effectively using FFT strategies with their children. During "practice," parents tried out FFT strategies onsite with their children and received immediate support and feedback from facilitators. To encourage parents to practice FFT strategies at home, facilitators: (a) provided parents with a hand-out summarizing the strategies at the end of each session; (b) sent reminder texts to parents every week; and (c) and asked parents to share their experiences practicing FFT strategies with other parents during the group sessions. A summary of FFT topics, activities, and strategies per session can be found in Table S1 (online supplementary material).

The active control condition involved one 90-min group session taking place in the school. The aim was to encourage parents to foster children's learning by playing simple games at home (e.g., puzzles, Legos®; inspired by activities used by Healey & Halperin, 2015). During the session, parents learned the games (60 min) and practiced the games with their children while receiving support and feedback (30 min). Parents received a handout. Offering this kind of active control condition helped increase school and family recruitment (compared with a business-as-usual control condition) and ruled out the possibility that any type of parenting session might have resulted in similar results as our strengths-based, culturally responsive program. Other RCTs have used a similar active control condition; for example, the effects of a professional program targeting pre-K and kindergarten teachers in Chile aiming to improve language and literacy skills were compared with those of an active condition, whereby teachers were provided books but no professional development (Yoshikawa et al., 2015).

School Recruitment and Randomization of Schools

We identified 35 Title 1 elementary schools (i.e., schools serving a significant percentage of students from low-income households) with

20% or higher of Latino students in one of the largest school districts in the U.S. located in the Southeast. Of the 35 schools invited to participate in a 3-year study in the fall of 2017, 17 schools accepted. Using a random number generator, we randomized schools to the treatment and control conditions in the spring of 2018. The randomization occurred at the school level to avoid potential spillover effects. Of the 17 schools that originally agreed to participate in the study, four schools (two in the treatment group and two in the control group) withdrew before the study started (early fall of 2018), either because they expressed feeling overwhelmed with other projects taking place at their school or because of principal turn-over. The final number of participating schools across the two years of the study was 13 schools. Of these 13 schools, one school participated in Year 1 but not Year 2 data collection. The COVID-19 pandemic prevented us from recruiting a third cohort of families and their children and from assessing our second cohort at the 5-month follow-up. In the Table S2, we present balance checks for the 13 (final sample) versus 17 schools (original sample) and discuss them in the Results section.

Statistical Power

Prior to data analysis, we performed a power analysis for the original 17 schools and three cohorts of children. The minimum detectable effect size (MDES) on the primary child outcome was .38, which was consistent with the overall effect size of home-based interventions (mean d = .47, range = .39–.55) reported by meta-analytic work (Manz et al., 2010). See Table S3 for full assumptions. Given COVID-19 disruptions and attrition, ultimately our MDES was .52 *SD* for our final posttest sample (13 schools, two cohorts), which was still within the range found by prior meta-analytic work (Manz et al., 2010), and .68 *SD* for our follow-up sample (13 schools, one cohort). Because our study was the second empirical study of FFT and the first RCT, it is comparable with an Institute of Education Science's Development and Innovation study (https://ies.ed.gov/). These types of studies are typically underpowered randomized studies that aim to evaluate whether the intervention merits larger-scale testing. In the Discussion section, we include power as a limitation.

Sample

We recruited 261 families over 2 years through an invitation letter sent to parents in their child's backpack and flyers distributed during the school's open house. We collected posttreatment narrative data from 234 children (on average, 18 children per school; M age = 67.17 months, SD = 4.08, 51% girls). There were 111 families in Year 1 (49 in treatment, 44.14%) and 123 families in Year 2 (37 in treatment, 30.08%). Of parents who completed at least some part of the demographic survey at pretest (n = 152; 65%), about 90% of parents were born outside of the U.S. and 22% had a GED diploma or higher. Families immigrated mostly from Central America (48%) and Mexico (41%).

Measures and Data Collection Procedure

We collected narratives from children at three time points: pretest (before program start; September), the end of the treatment posttest (1 to 2 weeks after program completion; November), and the 5-month follow-up posttest (April). For Cohort 1, we collected data at the three time points. For Cohort 2, we collected data at two time points (pretest and posttest), due to COVID-19 pandemic in the spring of 2020.

A team of 20 bilingual trained research assistants who were blinded to condition collected child data in schools. At each time point, children were individually pulled-out from the classroom by a bilingual research assistant and assessed in a separate room in the school. The total length of the assessment session was about 20 min and included the administration of the narrative task and other tasks assessing learning outcomes not included in this study (e.g., approaches to learning). The order in which the assessments was conducted was counterbalanced.

Narrative Task

A bilingual research assistant elicited two personal stories from the child: a story about a positive and a negative experience. Children were instructed to speak in the language in which they were most comfortable. Research assistants ensured that children knew they were bilingual and that they could speak in either language with them. The specific prompts (in English) were: "Tell me about a time that you were happy, one thing that you did that was really special and fun," "Tell me about a time that you got angry or mad at your sibling or friend." The order of presentation of the prompts was counterbalanced. Once the child started narrating, the research assistant encouraged continuation by echoing (repeating) what the child said or providing backchannel responses (e.g., uhmm, oh, wow!, uhuh, really?). If the child stopped talking, the research assistant prompted the child only once by saying "Anything else?" If the child declined to provide more information after this prompt, the narrative task was over. Bilingual research assistants were trained to avoid asking any other question and were instructed to stop if the narrative exceeded 5 min. However, no story reached that limit. The average duration of the narrative (including both positive and negative stories) was 2 min and 3 s (range = 58 s to 4 min and 21 s). This method of elicitation has been used by numerous studies and found to be successful at encouraging children to produce narratives in a seminaturalistic manner without imposing the research assistant's ideas or structure (Peterson & McCabe, 1983; Reese & Newcombe, 2007). Following prior research, we elicited personal emotional (positive/negative) experiences, as these tend to be more engaging to recount for Latino children than neutral events (Silva & McCabe, 1996). Narratives were audio recorded.

Covariates

We included two child-level covariates (age and sex) collected via parental consent form, four prerandom assignment school-level covariates taken from publicly available data from the North Carolina Department of Public Instruction (i.e., percent of students who participated in English language programs or special education programs, were Latino, were economically disadvantaged, and were retained in third grade), and two teacher-level covariates collected via teacher surveys at pretest (i.e., has master's degree and years of experience). We also collected parent demographic and home literacy data, but due to its missingness (rates ranged from 34% to 61%), these were not included in analyses.

In addition, in our primary specification we controlled for child's narrative length at pre- and posttest (measured as word tokens or total number of on-task, complete, and intelligible words produced by the

child during the narrative), following prior literature (Peterson et al., 1999; Reese & Newcombe, 2007). We present descriptive statistics for all features of child narrative language abilities by treatment status in Table 1. As shown, treatment children scored higher between pretest and the postintervention period on narrative coherence, literacy language features, word types, and narrative length, and scored the same on elaborations. Control group scores decreased between pre- and postintervention on all measures except for narrative coherence, which increased. In contrast, control group children improved on all measures between postintervention and the 5-month follow-up while treatment group children decreased slightly on all measures.

Cost Analysis Data Collection Procedure

The lead author of this article oversaw the implementation of FFT and in doing so, the author kept track of all expenditures for the program (i.e., personnel costs, materials). We used the Cost Out Tool (https://www.cbcse.org/costout) as described below for identifying nationally representative prices for personnel costs and space.

Coding

Child narratives were transcribed verbatim by a team of six trained, bilingual research assistants, with native to near-native oral and literate proficiency in both Spanish and English. The transcription followed conventions of the Child's Data Exchange System (CHILDES; MacWhinney, 2000); transcription occurred at the utterance level; utterances were identified as statements that were segmented by either closure, prolonged pausing, or intonation contour (Rowe, 2012). To check for consistency, 20% of narratives in each cohort were transcribed by more than one research assistant. The average interrater reliability of word-for-word (88%) and utterance transcription (86%) agreement across research assistants was high, calculated by dividing the total number of agreements by the sum of the total number of agreements and disagreements for each. The remaining 80% of narratives were independently transcribed; a research assistant reviewed the accuracy of these transcripts by listening to the audiotaped interaction, and correcting transcripts if there were discrepancies.

We coded children's narrative language abilities based on features assessed in prior RCTs: narrative coherence, elaborations, word types, literate language features (a measure of decontextualized talk), and narrative length. The same team of research assistants coded by hand for narrative coherence, elaborations, and literate language features. The team was first trained by a master coder and then participated in four rounds of reliability involving 8% of transcripts (n = 16; four transcripts per round). Interrater reliability was high (narrative coherence: κ = .90; elaborations: κ = .82, and literate language features: κ = .88, all ps $< .001$). After reaching reliability, an additional 25% of transcripts were double-coded by the team and disagreements were resolved by conferencing. Word types and narrative length were calculated automatically using the Child Language Analysis- CLAN program (MacWhinney, 2000). Children were given credit for every word/utterance, regardless of the language used (i.e., code-switching). Notably, the proportion of code-switching was low (M = 7.42% of utterances involved code-switching). Off-task and uncodable (unintelligible) child utterances were neither coded nor included in the analyses.

Narrative Coherence

We used a coding scheme at the story level (Reese et al., 2011) that assessed story structure using a 4-point scale (from 0 to 3) based on three dimensions: context, temporal, and theme coherence. Table S4 provides a full description of each of the scores (0 through 3) for each dimension and their corresponding examples. Here we provide anchor points. For context coherence, a score of 0 was assigned when no information about time or location was provided, and a score of 3 was assigned when both time and place were mentioned and were specific. For temporal, a score of 0 represented a narrative consisting of a list of actions with no temporal order and a score of 3 represented a narrative whereby 75% or more of relevant actions were temporally situated. For theme coherence, a score of 0 represented a narrative that was substantially off/topic, while a score of 3 represented a narrative that was on topic, substantially develops the theme and included

Table 1
Means (SD) for Child Assessments Across Cohorts 1 and 2

	Treatment (FFT)			Control		
Narrative features	Pretest	End of treatment (posttest)	5-month follow-up*	Pretest	End of treatment (posttest)	5-month follow-up*
Narrative coherence						
Mean (*SD*)	1.29 (0.66)	1.49 (0.66)	1.60 (0.59)	1.20 (0.62)	1.23 (0.67)	1.53 (0.56)
Min-Max	0 to 2.50	0.33 to 2.67	0.33 to 2.67	0.17 to 2.83	0.33 to 2.67	0.33 to 2.50
Elaborations						
Mean (*SD*)	6.63 (4.06)	6.24 (4.01)	5.79 (2.94)	5.89 (3.34)	4.40 (3.07)	6.73 (5.17)
Min-Max	1 to 23	1 to 19	1.50 to 15.50	1 to 16	1 to 15.50	1.50 to 25.50
Word types						
Mean (*SD*)	24.25 (13.63)	23.58 (13.67)	23.00 (11.91)	20.92 (11.55)	16.20 (10.27)	24.32 (15.19)
Min-Max	4 to 66	2 to 67.50	4.50 to 53.50	2 to 62	2 to 48.50	4.50 to 73.50
Literacy language features						
Mean (*SD*)	15.43 (12.38)	15.08 (12.08)	13.48 (9.48)	12.52 (9.44)	8.94 (7.53)	15.52 (14.30)
Min-Max	1 to 72	0 to 54	1 to 42	0 to 49	0 to 36.50	1.50 to 66.50

Note. Combined sample size for children with assessment data reported in this table was 198 for pretest; 234 for end-of-treatment; 100 for 5-month follow-up. Sample size was calculated based on the nonattriter children (n = 234) defined as those children with an end-of-treatment assessment.
* For the 5-month follow-up, only Cohort 1 data is available due to COVID-19.

a resolution or closure. The final score was the average across all three dimensions.

Elaborations

We used the coding scheme at the utterance level (Peterson et al., 1999; Reese & Newcombe, 2007), whereby elaborations are operationalized as the number of unique units of information about the past event produced by the child. For example, if the child said: "I went to the park," the child was given a score of 1 (one unit of information). If the child said: "I went to the park and played soccer," the child was given a score of 2 (two units of information). The final score was the sum of units of information produced by the child.

Word Types

Word types are the total number of unique words (i.e., uninflected word roots produced by the child; CLAN Manual, MacWhinney, 2000). Word types are an index of lexical diversity. The final score was the sum of all unique words (see Hammer & Sawyer, 2016 for a similar procedure).

Literate Language Features

Literate language features are a measure of decontextualized talk and a key discourse feature that develops during the kindergarten year (Westby, 1994). We used a coding scheme at the word level (Curenton & Justice, 2004) that focuses on four mutually exclusive literate language features related to the overall microstructure of the narrative: elaborated noun phrases (i.e., determiners, adjectives), adverbs, conjunctions (coordinating and subordinating conjunctions), and mental and linguistic terms (e.g., think, tell). Other coding schemes for decontextualized talk (e.g., number of utterances referencing past and future events, explanations, and pretend talk; Leech et al., 2018) were not appropriate because of the nature of the narrative task (i.e., the child was prompted to talk about past events). This coding scheme has been used in studies of racially minoritized (Curenton & Justice, 2004) and Latino populations (Schick, 2014). See Table S5 for a full description and examples of each category. The final score was the sum of all the literate language features produced by the child.

Children produced two personal stories (positive and negative past experience). If the child had complete data in both stories (positive/negative), the final score for each narrative feature (e.g., narrative coherence, elaborations) was the average score across stories (sum of positive and negative scores, divided by 2). If the child had data on one but not the other story (e.g., positive but not negative), the child's final score was the score in the story for which the child had complete data (e.g., positive). Overall, 189 children produced two stories and 45 produced one, with no difference by treatment status in one versus two. We compared this approach to using a listwise deletion approach and found similar results, which are available from authors upon request.

Data Analytic Plan

Missing Data

Missing data for student-level covariates included in analysis ranged from 0% to 15% (M = 12%; SD = 6%). At the teacher-level, 16% of teachers had missing data. From the original 261 children randomized into the study, we had end-of-treatment (posttest) data for 234 children (10% missing). For the 5-month follow-up, where no Cohort 2 data were collected because of COVID-19, we had complete data for 90% of Cohort 1. Missing data for posttest outcomes were the result of technical difficulties audiotaping the child's narrative, the child being absent the day of assessment, the child's family moving to another school or state. Based on Little's MCAR (χ^2 (df = 44) = 55.15, p = .12), data appear to have been missing at random. We used complete case as our primary approach.

Research Question 1

To estimate the impact of FFT, we first used OLS regressions to estimate an intent-to-treat (ITT) effect of being assigned to participate in the FFT program:

$$Y_{ics} = \beta_0 + \gamma(Treat)_s + \nu(pretest)_{ics} + X'_{ics} + \theta'_{cs} + \tau'_s + \varepsilon_{ics} \quad (1)$$

where Y is the child-level outcome of interest, i denotes child, c denotes classroom, and s denotes school. *Treat* is set to 1 if a given school was randomly assigned to treatment and 0 otherwise. We also included the pretest score for child i on outcome Y, child-level covariates (X'; child age and sex, child's narrative length at pretest and posttest, test language of pre- and posttest, and cohort), two characteristics of child i's kindergarten teacher (θ'; highest degree attained of teacher and teacher's years of experience), and several aggregate school-level covariates (τ'; percent of students who are Hispanic, economically disadvantaged, participating in English language programs or special education programs, and percent of students retained in third grade). We adjusted for clustering in schools within the treatment and control conditions using robust cluster-corrected standard errors at the school level.

Second, we used two-stage least squares regressions to estimate a treatment-on-the-treated (TOT) effect of being assigned to FFT and participating in at least one FFT session:

$$Attend_{ics} = \beta_0 + \gamma(Treat)_s + \nu(pretest)_{ics} + X'_{ics} + \theta'_{cs} + \tau'_s + \varepsilon_{ics} \quad (2)$$

$$Y_{ics} = \beta_0 + \gamma(\widehat{Attend})_{ics} + \nu(pretest)_{ics} + X'_{ics} + \theta'_{cs} + \tau'_s + \delta_{ics} \quad (3)$$

where assignment to FFT is used to predict attending at least one FFT session (Equation 2) and then this predicted value of attendance is used to estimate the effect of FFT (Equation 3). All other terms are defined as in Equation 1. In all, in the full sample, about 63% of treatment group members attended at least one session, while 0% of control families did, for a compliance rate of 63%.

Research Question 2

Our cost analysis work drew on leading advice in the field (Levin & McEwan, 2001) and several empirical examples (Jacob et al., 2016; Kabay et al., 2020). Specifically, to assess the costs of

FFT, we used a method developed by Levin (1975) in which all the ingredients that were used in implementing the program in our evaluation were described qualitatively and assigned a market price. We identified all implementation expenses in our trial and expenses based on requirements for replication, regardless of who paid for or contributed them (Levin & McEwan, 2001). Ingredients included bilingual personnel to lead the session, two assistants (one to provide child care and one to assist the lead in the session), training requirements, facilities to provide services and training, materials, FFT materials costs (i.e., folders, play food, play money, notepads; see Table S6 for all resources and prices), and transportation.

Next, we used the CostOut Tool (Center for Benefit-Cost Studies of Education, n.d.) to calculate nationally representative prices for personnel costs and space (in 2020 dollars). Materials were ordered online or purchased in large national chains; we assume these prices were already nationally representative in our analysis. We calculate total social costs both per school and per family (assuming 15 per school). For transparency, we also break costs out by what was spent in our trial (i.e., our direct costs) and then show nationally representative pricing, following the ingredients methods. Our national representative estimates represent cost to replicate the implementation.

There are a few assumptions in our nationally representative cost estimates that should be highlighted. First, in our trial, we hired lead and assistant bilingual facilitators through the first author's university as project staff. Were the program to be replicated, we assumed these roles would be filled by school staff, specifically an elementary school teacher and teacher assistants, at about 3.1% and .63 Full-Time Equivalent (FTE; used to convert part-time employee hours into full-time employee hours), respectively. We used national representative estimates for these roles and associated benefits from the Bureau of Labor Statistics included in the CostOut Tool (Center for Benefit-Cost Studies of Education, n.d.). Our FTE estimates come from our experience with this evaluation. We estimate that per school, the intervention requires approximately 50 hr of the lead facilitator's time (3 hr training, 10 hr recruitment, 4 1.5 hr sessions, 20 hr to purchase and organize materials, 1 hr for planning each session, 3 hr for sending text message to families, and 4 hr to debrief others in the school community). We estimate that each assistant would spend 10 hr total on the intervention (4 hr per session for prep and delivery, 2 hr of training total). Second, we assumed that the program would take place in an elementary school classroom. Although there was no direct charge to our team for space in our trial, we used the CostOut Tool to price out the opportunity cost of this space in our nationally representative estimates, assuming 1% usage per year. Third, we did not include transportation costs given that staff would already be at the schools as part of their regular jobs at the school. Finally, in our trial, there was no difference by site in costs. Accordingly, we assumed the same for our nationally representative estimates.

Data and study materials and analysis code are available upon request from the authors.

Results

Baseline Balance for Nonattriters

To test for baseline differences in child- and teacher-level characteristics of those assigned to treatment and control, we regressed each characteristic of interest on the treatment assignment variable, with a cluster correction for school where necessary. As shown in Table 2, we did not detect any statistically significant differences in child demographics, most of the pretest assessment scores, or teacher characteristics. There was a statistically significant difference on the literacy language features pretest such that the FFT group had higher scores than the control sample (*ES* = .27). We present balance checks on school-level characteristics in Table S1 (child- and teacher-level characteristics were not available for the four attriter schools) for both the final sample of schools ($N = 13$) and the original sample ($N = 17$). For either sample, we found no statistically significant differences on these characteristics, indicating evidence of balance by treatment status. Overall *F-test* of baseline equivalence (see Table 2), $F(10, 159) = 1.00$, $p = .45$, and *F-test* in the Table S1, $F(11, 1) = .21$, $p = .95$ were not statistically significant.

Attrition

At the child-level, 11.54% of children ($N = 27$) who were randomized were not assessed at either end-of-treatment or the 5-month follow-up. Differential attrition by treatment status was very minimal (1.37%); at the school level, we had a total attrition rate of 24% ($N = 4$ schools out of 17), with zero differential attrition by treatment status. According to What Works Clearinghouse conservative standards (IES, 2017), this evidence indicates low threat of bias.

Research Question 1: FFT Impacts

In Table 3, we present both the intent-to-treat and treatment-on-the-treated estimates for two models for narrative features, one of which includes child covariates only (M1) and one of which adds school and teacher covariates (M2). Given that results are generally stable across narrative features and across the two models, our preferred specification is the second model (Columns 7 and 10 for ITT and TOT respectively). We also show the results of the first-stage models predicting FFT attendance for the TOT models (Column 4).

FFT had statistically significant positive impacts on all narrative features measured at end-of-treatment. Children in schools randomly assigned to FFT (i.e., ITT) had more advanced narrative coherence, elaborations, word types, and literate language features relative to children in schools assigned to the control group. ITT effect sizes ranged from .42–.99 *SD* across the two ITT models. Similarly, children in the treatment group schools whose parents attended at least one FFT session (i.e., TOT) had higher scores on narrative coherence, narrative length, elaborations, word types, and literate language features compared with children in the control group schools. TOT effect sizes ranged from .44–1.76 *SD* across the two specifications. For Cohort 1, we also estimated both ITT and TOT effects at the 5-month follow-up. FFT did not have statistically significant impacts on children's elaborations, word types or literacy language features. However, we found positive and marginally significant impacts on narrative coherence for treatment-group children in schools assigned to FFT (effect sizes ranged from .23 to .35 *SD* across the two ITT models) and whose parents attended at least one of FFT session (effect sizes ranged

Table 2
Balance Checks on Nonattriters (n = 234 Children; n = 70 Teachers)

Variable	Treatment sample (n = 86) M (SD)	Control sample (n = 148) M (SD)	Raw difference	Effect size
Children demographics				
Gender	0.50	0.51	−0.01	−0.02
Age at pretest	66.98 (4.15)	67.28 (4.05)	−0.3	−0.07
Child-level assessment data				
Language of test at pretest is Spanish	0.50	0.59	−0.09	−0.18
Narrative coherence	1.29 (0.66)	1.20 (0.62)	0.10	0.16
Elaborations	6.63 (4.06)	5.89 (3.34)	0.74	0.20
Word types	24.25 (13.63)	20.92 (11.55)	3.33	0.27
Literacy language features	15.43 (12.38)	12.52 (9.44)	2.90*	0.27*
Narrative length (tokens)	39.55 (29.95)	32.43 (22.61)	7.12†	0.28†
Teacher-level data				
Teacher experience	9.21 (5.97)	11.52 (7.63)	−2.31	−0.33
Teacher has a BA (vs. a Master)	0.65	0.64	0.02	0.04

Note. Overall F-test is $F(10, 159) = 1.00$, $p = .45$. The overall F-test was run using all the variables in Table 1. Nonattriters are defined as those children who have a score at end-of-treatment (posttest). Out of the total 261 children randomized into the study, we did not have an end-of-treatment outcome data for 27 children. The raw difference column was obtained by regressing the characteristic of interest on intervention condition and clustering for school when applicable. Effect sizes were calculated by dividing the raw difference by the standard deviation of the full sample. Missing data for student-level covariates ranged from 0% to 15% (M = 12%; SD = 6%). At the teacher-level, 16% of teachers had missing data. Standard deviations are only reported for continuous variables.
† $p < .10$. * $p < .05$.

from .13 to .59 *SD*, $p < .10$ across the two TOT models) compared with children in the control condition.

Robustness Checks

Findings in our ITT and TOT models were not sensitive to alternative error structure choices (i.e., random intercepts for classroom and school) or other analytic decisions (e.g., inclusion vs. exclusion of covariates and inclusion of a set of dummy variables for the coder to account for the susceptibility to rater bias in narrative measures). Results of these analyses are available from authors upon request. We also compared results using complete case approach to replacing missing data with a constant and including a missing data indicator (IES, 2017). We found similar results across the data approaches (see Table S7).

Research Question 2: Program Costs

As shown in Table 4, in our trial, we spent $3,711 per school, or $247 per family with one child (in 2020 dollars). In nationally representative prices, following the ingredients approach (Levin & McEwan, 2001), replicating FFT would cost about $5,785 per school, or $386 per family with one child. Of these costs, $124 is in-kind (i.e., classroom space). These estimates represent the total social cost, which might be borne by a school using family engagement funds or perhaps a foundation or nonprofit working in partnership with the school. The two estimates also serve to bound costs. The largest expense in the nationally representative estimates was the teacher. If a community member were hired at a lower cost, the cost could be substantially lower, as our trial costs estimates help to illustrate.

Discussion

This RCT examined the impacts and costs of FFT, a strengths-based, culturally responsive intervention to improving kindergarteners' narrative language abilities by leveraging two valued sociocultural practices with built-in benefits: personal narratives and food routines. To assess the impacts of this intervention, we used a culturally grounded assessment (i.e., narrative task) that reflects how children integrate diverse language abilities that are foundational for reading comprehension (e.g., vocabulary, morphology) and leverages Latino children's ecocultural assets. Four features of narrative language abilities were measured: narrative coherence, elaborations, word types, and literate language features. We found positive impacts on all features of narrative language abilities at end-of-treatment and suggestive evidence of positive impacts on one feature (i.e., narrative coherence) at the 5-month follow-up. The costs of implementing this program were relatively low. Below we elaborate on each finding.

The effect sizes of FFT impacts on narrative language abilities at the end-of-treatment were moderate using ITT models (d = .69–.99) and large using TOT models (d = 1.21–1.76), the latter of which considers not only whether families were assigned to the intervention condition, but also whether they attended intervention sessions. The effect sizes at the 5-month follow-up on one feature of narrative language abilities was small using ITT models (d = .35) and moderate using TOT models (d = .59). Our findings are aligned with those of a recent meta-analysis of school-based interventions focused on language and targeting children in preschool, elementary, and secondary school (Rogde et al., 2019). While small effects on vocabulary and grammatical knowledge were found at end-of-treatment (gs ranged from .17 to .19), moderate effects on narrative language abilities were found (g = .42).

Our findings complement prior work (Leyva et al., 2021), whereby positive moderate effects of FFT were observed on a

Table 3
FFT Impacts on Child Narrative Outcomes

	Model 1					Model 2				
	ITT		TOT			ITT		TOT		
Narrative features	Estimate	ES	First stage estimate	Estimate	ES	Estimate	ES	First stage estimate	Estimate	ES
				End of treatment						
Narrative coherence	0.28*	0.42	0.64***	0.44*	0.66	0.46***	0.69	0.57***	0.81***	1.21
	(0.12)			(0.20)		(0.12)			(0.27)	
Elaborations	1.58*	0.52	0.64***	2.49*	0.81	2.43**	0.79	0.57***	4.28**	1.40
	(0.64)			(1.06)		(0.77)			(1.52)	
Word types	5.93**	0.58	0.64***	9.24**	0.90	8.45***	0.82	0.57***	14.69***	1.43
	(1.92)			(3.19)		(2.37)			(4.78)	
Literacy language features	5.00**	0.66	0.64***	7.87**	1.05	7.49**	0.99	0.57***	13.25**	1.76
	(1.74)			(3.00)		(2.37)			(4.93)	
				5-month follow-up (Cohort 1 only)						
Narrative coherence	0.14	0.23	0.56***	0.08	0.13	0.21†	0.35	0.59***	0.36†	0.59
	(0.17)			(0.10)		(0.12)			(0.20)	
Elaborations	−0.84	−0.17	0.56***	−0.47	−0.09	0.48	0.10	0.59***	0.82	0.17
	(1.32)			(0.75)		(1.09)			(1.67)	
Word types	−1.67	−0.11	0.56***	−0.93	−0.06	1.39	0.09	0.59***	2.37	0.16
	(4.37)			(2.53)		(3.67)			(5.60)	
Literacy language features	−2.77	−0.20	0.56***	−1.54	−0.11	0.56	0.04	0.59***	0.95	0.07
	(3.84)			(2.19)		(3.17)			(4.88)	

Note. Standard errors in parentheses. Effect sizes (ES) are standardized on the standard deviation of the control group mean. Model 1 includes controls for pretest language, posttest language, child age, child gender, pretest narrative length, posttest narrative length, and an indicator for cohort (for Post 1 outcomes only). Model 2 adds school-level covariates (% Hispanic, % Limited English Proficient, % special education, % economically disadvantaged) and teacher-level covariates (has master's degree, years of experience). We defined compliers as parents who attended at least one FFT meeting. Sample sizes $N = 191$ on Post 1 outcomes and $N = 90$ on Post 2 outcomes (Cohort 1 only). ITT = intent-to-treat; TOT = treatment-on-the-treated.
† $p < .10$. * $p < .05$. ** $p < .01$. *** $p < .001$.

nonstandardized measure of expressive vocabulary ($d = .33$ using ITT and .57 using TOT models), but no effects were found on standardized measures of language and literacy. Taken together, these findings suggest that is critical for RCTs to include both standardized and nonstandardized assessments of children's learning outcomes to better capture improvements (Slavin, 2019), and more importantly, to use culturally grounded measures that avoid assuming certain knowledge or experience, particularly when assessing racially minoritized and linguistically diverse children (Fiestas & Pena, 2004). Currently, there is wide variability in the way narratives are elicited and measured by studies which prevents us from directly comparing the gains in narrative language abilities documented here with those reported by previous work with Latino kindergarteners (Miller et al., 2006; Uccelli & Paez, 2007). This variability includes differences in whether children are required to tell or retell a story, whether a wordless picture book is used or not, and how narrative features are scored. Hence, culturally grounded measures such as narrative tasks deserve more systematization in their elicitation and measurement procedures and rigorous psychometric attention and validation so they can be readily adopted by other RCTs.

Positive impacts on all four features of narrative language abilities at end-of-treatment align with the explicit focus in FFT on personal narratives during food routines. Furthermore, children's language abilities were assessed within the context of narratives, a context that Latino children are familiar with. Our narrative assessments likely were more sensitive to detecting changes in children's language development than other (standardized) language tests because they are closer in context to the targeted instruction. Others have also noted that measures aligned with the content of the intervention are more sensitive to treatment effects than independent standardized tests (Slavin, 2019).

Our findings also resonate with those of recent meta-analytic work on language comprehension interventions in elementary schools in the U.S., which concluded that interventions that bring together several components of language (e.g., vocabulary, morphology, syntax) and provide children with authentic opportunities to use language for expression and comprehension, like FFT does via promotion of personal narratives during family food routines, are more likely to be effective than those that focus on a single component or provide little opportunities for authentic use of language expression and comprehension skills (Silverman et al., 2021).

We found suggestive evidence that the positive impacts of FFT persisted through the 5-month follow-up in one feature of narrative language abilities: narrative coherence. There are two plausible explanations for this marginally significant finding. First, we might have been underpowered to detect statistically significant impacts at this time point. Due to COVID-19 pandemic, we were only able to assess one cohort, not the full sample, at the 5-month follow-up. Second, impacts on narratives features are sometimes detected only at the 1-year follow-up (Peterson et al., 1999). These sleeper effects or long-term gains in narrative features might be a result of parents gradually, rather than instantly, adopting strategies to effectively engage their children in conversations (Peterson et al., 1999; Thierry & Sparks, 2013).

Nevertheless, this suggestive evidence seems promising. Narrative coherence is children's ability to tell a story that makes sense

Table 4
Food For Thought (FFT) Costs (in 2020 Dollars)

Ingredients	Spent in present trial	Nationally representative prices, with ingredients method approach
Lead bilingual facilitator	$1,355	$2,907
Bilingual facilitator assistants (2 assistants)	$202	$875
Materials	$1,880	$1,880
Travel	$274	$0
Space	$0	$124
Total cost per school	$3,711	$5,786
Total cost per family with one child	$247	$386

Note. FFT materials costs are detailed in Table S6. We used the CostOut Tool (Center for Benefit-Cost Studies of Education, n.d.) to calculate nationally representative prices for personnel costs and space (in 2020 dollars). Materials were ordered online or purchased in large national chains; we assume these prices were already nationally representative in our analysis. As explained in the text, for our trial we hired lead and assistant bilingual facilitators through the first author's university as project staff. For nationally representative estimates, we assumed these roles would be filled by school staff, specifically an elementary school teacher and teacher assistants, at about 3.1% and .63 FTE, respectively. We used national representative estimates for these roles and associated benefits from the Bureau of Labor Statistics included in the CostOut Tool (Center for Benefit-Cost Studies of Education, n.d.). We estimate that per school, the intervention requires approximately 50 hr of the lead facilitator's time (3 hr training, 10 hr recruitment, 4 1.5 hr sessions, 20 hr to purchase and organize materials, 1 hr for planning each session, 3 hr for sending text message to families, and 4 hr to debrief others in the school community). We estimate that each assistant would spend 10 hr total on the intervention (4 hr per session for prep and delivery, 2 hr of training total).

to a naïve listener in terms of space, time, and theme. Planning and organizing the discourse and monitoring for coherence are critical to language development (McCabe & Rollins, 1994; Reese et al., 2010). Others have found that children's narrative coherence positively relates to their expressive language (Reese et al., 2011), and uniquely predicts their reading proficiency (Miller et al., 2006). Thus, it is encouraging that this one narrative feature was positively impacted by FFT beyond the immediate posttest.

We also found that the costs of implementing FFT are low, particularly when compared with other literacy programs implemented in schools such as the Reading Partners program ($710 per student; Jacob et al., 2016) and Sound Partner program ($791 per student; Hollands et al., 2016). To be sure, the targets and benefits of these programs are different than FFT's. But given the lack of cost data on family-targeted programs like FFT, this cost comparison does provide at least a touch point on feasibility. FFT is likely within reach of districts seeking additional evidence-based approaches to promoting literacy for Latino kindergarteners, especially given that under the Every Student Succeeds Act (U.S. Department of Education, 2015), school districts are required to allocate at least 1% of federal funds for parental engagement activities in schools serving students from low-income households.

Limitations and Future Directions

Our study has several important limitations. We collected brief speech samples from children, although the length is developmentally appropriate for children of this age (see MacWhinney, 2000). In future studies, using home audio recording devices (e.g., LENA®) to collect longer speech samples might better capture how children integrate diverse oral language skills (e.g., vocabulary, morphology, syntax) and other cognitive skills (e.g., planning and monitoring story coherence) in the context of telling stories (Florit & Cain, 2011; McCabe & Rollins, 1994). Note, however, that LENA audio still requires manual transcription and coding to produce parent and child language measures beyond word count and conversational turns. We elicited personal emotional (positive/negative) experiences as part of our culturally grounded assessment (i.e., narrative task), given prior work showing that Latino children are more eager to share these experiences (Silva & McCabe, 1996). It is uncertain whether similar findings would be detected had we used other prompts to elicit personal narratives. For example, asking children to plan for an upcoming event such as grocery shopping or going to the beach (Hudson et al., 1995). We focused on four of the most common narrative features analyzed by prior RCTs of interventions parent–child personal narratives (narrative coherence, elaborations, word types, and literate language features) and used narrative length as a covariate following prior research (Peterson et al., 1999; Reese & Newcombe, 2007; Reese et al., 2010). We acknowledge that there are other narrative features such syntactic complexity (Rowe, 2012) that might be important to analyze in future work. In addition, ours is a relatively small RCT. Further replication with larger sample is needed in the next phase of development and evidence building for FFT. Finally, as noted in our Cost Analysis section, we made several assumptions in our nationally representative cost analysis —that FFT would be implemented by elementary school staff, take place in an elementary school, and that there would be no difference in costs by site. These assumptions may not hold in a real-world replication of the program and costs would differ accordingly.

Developmental and Educational Implications

Our findings have important implications for Latino kindergarteners' language and literacy skills and their families. Several longitudinal studies have demonstrated that oral language abilities (including narrative skills) develop before formal reading instruction begins and are an important precursor to reading comprehension success (for a review see Rogde et al., 2019). Thus, promoting

such oral language abilities, like FFT does, has the potential to make an important impact in the literacy development of Latino children, which is critical to their academic success (Uccelli et al., 2019).

More than a decade ago, Common Core State Standards advocated for a strong focus on the development of oral language skills (National Governors Association Center for Best Practices & Council of Chief State School Officers, 2010). In recent years, the debate over the "science of reading" has received great attention, with an increasing emphasis on decoding and teaching explicit phonics. However, as many have argued, both decoding and oral language abilities (including narrative skills) are equally important and emphasizing only one of them can be detrimental to the development of children's literacy competence (Silverman et al., 2021). Both skills can be supported through direct instruction in the classroom but also at home, through authentic opportunities. In fact, home-based language-meaning activities such as storytelling predict larger vocabulary gains over a year than activities which target decoding skills, suggesting that it might be beneficial for families to engage in oral language activities as a complement to the decoding/written language activities that are typically the focus of school curricula (McCormick et al., 2020). FFT provides an example of how leveraging personal narratives during food routines can improve children's narrative language abilities, a foundational skill to reading comprehension.

References

Ansari, A. (2017). The selection of preschool for immigrant and native-born Latino families in the United States. *Early Childhood Research Quarterly*, *41*(4), 149–160. https://doi.org/10.1016/j.ecresq.2017.07.002

Center for Benefit-Cost Studies of Education (n.d.). CostOut Tool. Retrieved from https://www.cbcse.org/costout

Curenton, S. M., & Justice, L. M. (2004). African American and Caucasian preschoolers' use of decontextualized language: Literate language features in oral narratives. *Language, Speech, and Hearing Services in Schools*, *35*(3), 240–253. https://doi.org/10.1044/0161-1461(2004/023)

Dickinson, D. K., & Tabors, P. O. (2001). *Beginning literacy with language: Young children learning at home and school*. Paul H Brookes Publishing.

Dowdall, N., Melendez-Torres, G. J., Murray, L., Gardner, F., Hartford, L., & Cooper, P. J. (2020). Shared picture book reading interventions for child language development: A systemic review and meta-analysis. *Child Development*, *91*(2), e383–e399. https://doi.org/10.1111/cdev.13225

Eisenberg, A. (2002). Maternal teaching talk within families of Mexican descent: Influences of task and socioeconomic status. *Hispanic Journal of Behavioral Sciences*, *24*(2), 206–224. https://doi.org/10.1177/0739986302024002006

Evans, A., Chow, S., Jennings, R., Dave, J., Scoblick, K., Sterba, K. R., & Loyo, J. (2011). Traditional foods and practices of Spanish-speaking Latina mothers influence the home food environment: Implications for future interventions. *Journal of the American Dietetic Association*, *111*(7), 1031–1038. https://doi.org/10.1016/j.jada.2011.04.007

Fiestas, E., & Pena, C. (2004). Narrative discourse in bilingual children: Language and task effects. *Language, Speech, and Hearing Services in Schools*, *35*(2), 155–168. https://doi.org/10.1044/0161-1461(2004/016)

Florit, E., & Cain, K. (2011). The simple view of reading: Is it valid for different types of alphabetic orthographies? *Educational Psychology Review*, *23*(4), 553–576. https://doi.org/10.1007/s10648-011-9175-6

Garcia-Coll, C., Lamberty, G., Jenkins, R., McAdoo, H. P., Crnic, K., Wasik, B. H., & Vázquez García, H. (1996). An integrative model for the study of developmental competencies in minority children. *Child Development*, *67*(5), 1891–1914. https://doi.org/10.2307/1131600

Hammer, C. S., & Sawyer, B. (2016). Effects of a culturally responsive interactive book-reading intervention on the language abilities of preschool dual language learners: A pilot study. *Dialog*, *18*, 59–79.

Healey, D. M., & Halperin, J. M. (2015). Enhancing Neurobehavioral Gains with the Aid of Games and Exercise (ENGAGE): Initial open trial of a novel early intervention fostering the development of preschoolers' self-regulation. *Child Neuropsychology*, *21*(4), 465–480. https://doi.org/10.1080/09297049.2014.906567

Hollands, F. M., Kieffer, M. J., Shand, R., Pan, Y., Cheng, H., & Levin, H. M. (2016). Cost-effectiveness analysis of early reading programs: A demonstration with recommendations for future research. *Journal of Research on Educational Effectiveness*, *9*(1), 30–53. https://doi.org/10.1080/19345747.2015.1055639

Hudson, J. A., Shapiro, L. R., & Sosa, B. B. (1995). Planning in the real world: Preschool children's scripts and plans for familiar events. *Child Development*, *66*(4), 984–998. https://doi.org/10.2307/1131793

Institute of Educational Sciences. (2017). *What Works Clearinghouse Practice Guides*. https://ies.ed.gov/ncee/wwc/practiceguides

Institute of Educational Sciences (IES). (2020). *IES Cost Analysis Start Kit, Version 1.0*. https://ies.ed.gov/seer/pdf/IES_Cost_Analysis_Starter_Kit_V1.pdf

Jacob, R., Armstrong, C., Bowden, A., & Pan, Y. (2016). Leveraging volunteers: An experimental evaluation of a tutoring program for struggling readers. *Journal of Research on Educational Effectiveness*, *9*(sup1), 67–92. https://doi.org/10.1080/19345747.2016.1138560

Kabay, S., Weiland, C., & Yoshikawa, H. (2020). Costs of the Boston public prekindergarten program. *Journal of Research on Educational Effectiveness*, *13*(4), 574–600. https://doi.org/10.1080/19345747.2020.1799463

Kermani, H., & Janes, H. A. (1999). Adjustment across task in maternal scaffolding in low-income Latino immigrant families. *Hispanic Journal of Behavioral Sciences*, *21*(2), 134–153. https://doi.org/10.1177/0739986399212002

Leech, K., Wei, R., Harring, J. R., & Rowe, M. L. (2018). A brief parent-focused intervention to improve preschoolers' conversational skills and school readiness. *Developmental Psychology*, *54*(1), 15–28. https://doi.org/10.1037/dev0000411

Levin, H. (1975). Cost-effectiveness analysis in evaluation research. In M. Guttentag & E. L. Struening (Eds.), *Handbook of evaluation research* (Vol. 2, pp. 89–122). Sage.

Levin, H., & McEwan, P. (2001). *Cost-effectiveness analysis: Methods and applications* (2nd ed.). Sage.

Leyva, D., & Skorb, L. (2017). Food For Thought: Family food routines and literacy in Latino kindergarteners. *Journal of Applied Developmental Psychology*, *52*, 80–90. https://doi.org/10.1016/j.appdev.2017.07.001

Leyva, D., Weiland, C., & Shapiro, A. (2018). As predicted: Evaluation of the effectiveness of the Food For Thought program (#13229). https://aspredicted.org/blind.php?x=62m22i

Leyva, D., Weiland, C., Shapiro, A., Yeomans-Maldonado, G., & Febles, A. (2021). A Strengths-Based, Culturally Responsive Family Intervention Improves Latino Kindergarteners' Vocabulary and Approaches to Learning. *Child Development*. https://doi.org/10.1111/cdev.13698

MacWhinney, B. (2000). *The CHILDES project: tools for analyzing talk*. Erlbaum.

Manz, P. H., Hughes, C., Barnabas, E., Bracaliello, C., & Ginsburg-Block, M. (2010). A descriptive review and meta-analysis of family-based emergent literacy interventions: To what extent is the research applicable to low-income, ethnic-minority or linguistically-diverse young children? *Early Childhood Research Quarterly*, *25*(4), 409–431. https://doi.org/10.1016/j.ecresq.2010.03.002

McCabe, A., & Rollins, P. (1994). Assessment of preschool narrative skills. *American Journal of Speech-Language Pathology*, *3*(1), 45–56. https://doi.org/10.1044/1058-0360.0301.45

McCormick, M. P., Weissman, A. K., Weiland, C., Hsueh, J., Sachs, J., & Snow, C. (2020). Time well spent: Home learning activities and gains in

children's academic skills in the prekindergarten year. *Developmental Psychology*, *56*(4), 710–726. https://doi.org/10.1037/dev0000891

Melzi, G., Schick, A. R., & Kennedy, J. L. (2011). Narrative elaboration and participation: Two dimensions of maternal elicitation style. *Child Development*, *82*(4), 1282–1296. https://doi.org/10.1111/j.1467-8624.2011.01600.x

Melzi, G., Schick, A., & Scarola, L. (2019). Literacy interventions that promote home-to-school links for ethnoculturally diverse families of young children. In C. McWayne, F. Doucet, & S. Sheridan (Eds.), *Ethnocultural diversity and the home-to-school link* (pp. 123–143). Springer. https://doi.org/10.1007/978-3-030-14957-4_8

Michie, S., Jochelson, K., Markham, W. A., & Bridle, C. (2009). Low-income groups and behaviour change interventions: A review of intervention content, effectiveness and theoretical frameworks. *Journal of Epidemiology and Community Health*, *63*(8), 610–622. https://doi.org/10.1136/jech.2008.078725

Miller, J., Heilmann, J., Nockerts, A., Iglesias, A., Fabiano, L., & Francis, D. (2006). Oral language and reading in bilingual children. *Learning Disabilities Research & Practice*, *21*(1), 30–43. https://doi.org/10.1111/j.1540-5826.2006.00205.x

Murphey, D., Guzman, L., & Torres, A. (2014). *America's Hispanic children: Gaining ground, looking forward. Publication #2014-38. Child Trends*. http://www.childtrends.org

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common Core State Standards for English language arts*. National Governors Association Center for Best Practices, Council of Chief State School Officers.

Perez-Brena, N. J., Rivas-Drake, D., Toomey, R. B., & Umaña-Taylor, A. J. (2018). Contributions of the integrative model for the study of developmental competencies in minority children: What have we learned about adaptive culture? *American Psychologist*, *73*(6), 713–726. https://doi.org/10.1037/amp0000292

Peterson, C., & McCabe, A. (1983). *Developmental psycholinguistics: Three ways of looking at a child's narrative*. Plenum Press. https://doi.org/10.1007/978-1-4757-0608-6

Peterson, C., Jesso, B., & McCabe, A. (1999). Encouraging narratives in preschoolers: An intervention study. *Journal of Child Language*, *26*(1), 49–67. https://doi.org/10.1017/S0305000998003651

Reese, E., & Newcombe, R. (2007). Training mothers in elaborative reminiscing enhances children's autobiographical memory and narrative. *Child Development*, *78*(4), 1153–1170. https://doi.org/10.1111/j.1467-8624.2007.01058.x

Reese, E., Haden, C. A., Baker-Ward, L., Bauer, P., Fivush, R., & Ornstein, P. A. (2011). Coherence of personal narratives across the lifespan: A multidimensional model and coding method. *Journal of Cognition and Development*, *12*(4), 424–462. https://doi.org/10.1080/15248372.2011.587854

Reese, E., Leyva, D., Sparks, A., & Grolnick, W. (2010). Maternal elaborative reminiscing increases low-income children's narrative skills relative to dialogic reading. *Early Education and Development*, *21*(3), 318–342. https://doi.org/10.1080/10409289.2010.481552

Rogde, K., Hagen, A., Melby-Lervag, M., & Lervag, A. (2019). The effect of linguistic comprehension instruction on generalized language and reading comprehension skills: A systematic review. *Campbell Systematic Reviews*, *15*(4), e1059. https://doi.org/10.1002/cl2.1059

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development*, *83*(5), 1762–1774. https://doi.org/10.1111/j.1467-8624.2012.01805.x

Rowe, M. L. (2013). Decontextualized language input and preschoolers' vocabulary development. *Seminars in Speech and Language*, *34*(4), 260–266. https://doi.org/10.1055/s-0033-1353444

Save the Children. (2017). *International Development and Early Learning Assessment (IDELA). Overview of assessment*. https://idela-network.org/community-of-practice/topic/how-to-calculate-scores-for-the-four-idela-domains/

Schick, A. (2014). Home-school literacy experiences of Latino preschoolers: Does continuity predict positive child outcomes? *Journal of Applied Developmental Psychology*, *35*(4), 370–380. https://doi.org/10.1016/j.appdev.2014.05.006

Silva, M., & McCabe, A. (1996). Vignettes of the continuous and family ties: Some Latino American traditions. In A. McCabe (Ed.), *Chameleon readers: Teaching children to appreciate all kinds of good stories* (pp. 116–136). McGraw Hill.

Silverman, R., Johnson, E., Keane, K., & Khanna, S. (2021). Beyond decoding: A meta-analysis of the effects of language comprehension interventions on K-5 students' language and literacy outcomes. *Reading Research Quarterly*, *55*, S207–S233. https://doi.org/10.1002/rrq.346

Slavin, R. (2019). *Developer- and research-made measures*. https://robertslavinsblog.wordpress.com/2019/10/24/developer-and-researcher-made-measures/

Thierry, K., & Sparks, A. (2013). Latino families engage in elaborative conversations: Effects on children's recall and vocabulary. *Journal of Child and Family Studies*, *28*, 2297–2311. https://doi.org/10.1007/s10826-019-01446-0

Uccelli, P., & Paez, M. M. (2007). Narrative and vocabulary development of bilingual children from kindergarten to first grade: Developmental changes and associations among English and Spanish skills. *Language, Speech, and Hearing Services in Schools*, *38*(3), 225–236. https://doi.org/10.1044/0161-1461(2007/024)

Uccelli, P., Demir-Lira, Ö. E., Rowe, M. L., Levine, S., & Goldin-Meadow, S. (2019). Children's early decontextualized talk predicts academic language proficiency in midadolescence. *Child Development*, *90*(5), 1650–1663. https://doi.org/10.1111/cdev.13034

U.S. Census Bureau. (2019*). Quick Facts. Hispanic or Latino. Population Estimates July 1, 2019*. https://www.census.gov/quickfacts/fact/table/U.S.#

U.S. Department of Education. (2015). *Every Student Succeeds Act*. https://www.ed.gov/essa?src=rn

Westby, C. (1994). The effects of culture on genre, structure, and style of oral and written texts. In G. Wallach & K. Butler (Eds.), *Language learning disabilities in school-age children and adolescents* (pp. 180–218). MacMillan.

Woodcock, R. W., Muñoz-Sandoval, A. F., Ruef, M., & Alvarado, C. G. (2005). *Woodcock Muñoz Language Survey—Revised*. Riverside Publishing.

Yoshikawa, H., Leyva, D., Snow, C. E., Treviño, E., Barata, M. C., Weiland, C., Gomez, C. J., Moreno, L., Rolla, A., D'Sa, N., & Arbour, M. C. (2015). Experimental impacts of a teacher professional development program in Chile on preschool classroom quality and child outcomes. *Developmental Psychology*, *51*(3), 309–322. https://doi.org/10.1037/a0038785

Received September 16, 2021
Revision received November 27, 2021
Accepted December 10, 2021 ■