Pre-Analysis Plan: Filming Interactions to Nurture Development in Rural China

Yue Ma and Shanshan Li

August 2026

Introduction

Responsive caregiver-child interaction is an important input into early language, cognitive, and socioemotional development, yet caregivers may face multiple barriers to providing responsive care (Cunha et al., 2026). They may not fully understand which interactions support development, may have difficulty recognizing and responding to child cues, may lack confidence or motivation, may perceive the caregiver-child relationship less positively, or may face psychological stress that makes responsive engagement more difficult (Cunha et al., 2026).

This study examines whether Filming Interactions to Nurture Development (FIND), a 10-week strengths-based, personalized video-feedback parenting intervention, can improve responsive caregiving among families in rural China and examines caregiver-level intermediate outcomes consistent with the channels through which these behavioral changes may occur. We have three specific objectives. First, we seek to determine whether FIND improves responsive caregiver-child interaction and parenting investment/developmental stimulation. Second, we seek to test four theoretically motivated intermediate channels: (1) parenting beliefs; (2) parenting self-efficacy; (3) caregiver-child relationship; and (4) psychological capacity. Third, we explore secondary outcomes, including broader parenting practices, caregiver well-being, and child development. Because the endpoint follows a 10-week intervention, our main short-run hypotheses concern caregiver intermediate outcomes and behavior rather than large immediate changes in child development.

Research Design

We conduct a cluster-randomized controlled trial in 105 rural villages across four counties in Hubei Province, China. Villages are eligible if they: (1) have at least eight resident children aged 12–30 months; (2) are located outside minority townships where the majority population is non-Han; and (3) have at least one village resident willing to participate, as recommended by the village committee. From the pool of eligible villages, 105 villages are randomly selected. To minimize spillovers, neighboring villages are replaced so that sampled villages are at least 2 km apart. The target population consists of children aged 12–30 months at baseline and their primary caregivers, primarily mothers and grandmothers. Eligible households are invited to participate provided they do not plan to move during the study period.

Randomization occurs at the village level. We use a stratified design at the county level and randomly assign villages with a 1:1 allocation ratio to the FIND intervention (52 villages) or a pure control group

(53 villages). The enrolled sample comprises 827 child-caregiver pairs, or approximately eight pairs per village. The program-delivery window runs from May 31, 2026 through September 3, 2026. The registration record will document the timing of PAP finalization relative to treatment assignment, endline data access, and outcome analysis.

The intervention is a culturally adapted FIND program translated and localized for rural China. FIND uses positive reinforcement to identify and highlight naturally occurring, developmentally supportive caregiver-child interactions. It is grounded in the serve-and-return framework: a child "serves" by initiating interaction through words, gestures, actions, or attention, and a caregiver "returns" by noticing and responding. The program introduces five elements of serve and return: Sharing the Child's Focus; Supporting and Encouraging; Naming; Back-and-Forth Interaction; and Endings and Beginnings (Balldin et al., 2018; Liu et al., 2025). The control group does not receive FIND during the study period and participates in the same research data-collection procedures.

The intervention has five broader design components that motivate the theory of change.

- First, education and coaching on the five FIND elements provide the substantive content of responsive caregiving.
- Second, feedback is personalized using video of the caregiver interacting with their own child.
- Third, FIND uses a strengths-based approach that reinforces successful behaviors already occurring.
- Fourth, trained parenting coaches (from the local community) deliver the program through home visits.
- Fifth, the program provides 10 repeated home-based sessions, allowing caregivers to practice, receive feedback, and reinforce responsive behaviors over time. Caregiver-child interactions are filmed and feedback is provided using personalized edited films.

FIND Project Timeline

Project stage	Proposed timing	Main activities
Intervention adaptation	September 2024 to February 2026	Localize FIND materials; finalize coach and editor manuals; adapt videos, examples, and delivery procedures for the rural Chinese context
Pilot	May to August 2025	Pilot the intervention and research instruments; test coaching, video editing, survey and data-management procedures; revise protocols
Baseline survey	May 10, 2026	Recruit 827 families; obtain informed consent; collect baseline caregiver, child, and household data
Intervention stage	May to September 2026	Randomization; deliver FIND sessions; monitor attendance, implementation fidelity, coaching quality, and intervention completion
Endline survey	September 03, 2026	Survey all participating families; collect caregiver-child videos and LENA recordings from the intensive subsample

Power Analysis

Ex ante power calculations were based on 100 randomized villages with approximately eight child-caregiver pairs per village; the realized sample includes 105 villages and 827 child-caregiver pairs. The calculations assume cluster-level randomization, a two-sided 5 percent significance level, and an intra-cluster correlation of 0.01. Reference effect sizes are drawn from prior parenting-intervention studies. For responsive parenting behavior, an effect size of 0.39 standard deviations implies power of approximately 0.99 (Emmers et al., 2021). For the home learning environment, an effect size of 0.31 standard deviations implies power of approximately 0.99 (Cunha et al., 2024). For child development, an effect size of 0.26 standard deviations implies approximately 0.94 power (Emmers et al., 2021).

Balance Checks

We assess balance on prespecified baseline characteristics across the FIND and control groups. Child characteristics include age, gender, birth prematurity, firstborn status, and developmental outcomes; caregiver and household characteristics include whether the mother is the primary caregiver, primary caregiver age, primary caregiver education, primary caregiver self-efficacy, caregiver-child relationship, stimulation behaviors, and household asset index. For each baseline characteristic, we regress the characteristic on treatment assignment and include county-level randomization-strata fixed effects. We cluster standard errors at the village level. We also report a joint test of the prespecified baseline covariates. Balance tests are descriptive and do not determine whether covariates are included in the primary specification.

Outcomes and Measurement

The outcome structure in this pre-analysis plan follows the study's theory of change and measurement framework. Measures that are behavioral manifestations of responsiveness are kept separate from intermediate outcomes hypothesized to precede those behaviors. A supplementary outcome codebook finalized before endline-outcome analysis will list the exact item set, scoring rule, direction, and missing-component rule for every primary and intermediate measure.

We prespecify two main behavioral outcome families.

- The first family is responsive caregiving / reciprocal interaction.
    - Self-reported responsive skills and the responsiveness subscale of the Responsive Caregiving Rating Scale for Children Aged 0 to 4 Years (RCRS).
    - Observed measures:
        - Video-coded serve-and-return responsiveness, including cue-contingent caregiver responsiveness, coded from a 10–15-minute recorded caregiver-child interaction.
        - LENA conversational turn counts and a measure of child-initiated conversational turn counts.
        - LENA and video outcomes will be analyzed in their prespecified measurement subsamples. Because collecting and processing day-long audio recordings and

caregiver-child interaction videos is logistically intensive, these outcomes will not necessarily be available for the full randomized sample.

- The second primary behavioral outcome family is parenting investment and developmental stimulation.
    - Measures in this family include the Family Care Indicators (FCI).
    - Additional measures include the Promoting Cognitive and Emotional Development subscale of the RCRS and prespecified caregiver-child interaction questions. This family captures the frequency and breadth of developmentally supportive activities and is conceptually distinct from reciprocal responsiveness.

We prespecify four intermediate outcome domains.

- Parenting beliefs
    - Beliefs about responsive parenting investment
    - Beliefs about the returns to broader parenting investment
    - Scenario-based parenting-belief measures
- Parenting self-efficacy will be measured using the short form of the Tool to Measure Parenting Self-Efficacy (TOPSE) and the Chinese Parenting Sense of Competence Scale (C-PSOC)
- Caregiver-child relationship will be measured primarily using the Mothers' Object Relations Scales (MORS), adapted as needed for mothers, grandmothers, and other primary caregivers.
- Psychological capacity will be measured primarily using the Parental Stress Scale.

We will also examine secondary and exploratory caregiver-level outcomes: parenting discipline style, perceived social support, screen-related parenting and exposure, and broader caregiver well-being, including DASS-21 and happiness/life satisfaction. Downstream child-development outcomes include CDI language/vocabulary, CREDI development, and ASQ-SE socioemotional development. Any additional child assessments not explicitly listed in the outcome codebook will be labeled exploratory. Because these are distal outcomes at an immediate post-intervention endpoint, the FIND theory of change does not require large short-run child-development effects.

Empirical Specifications

Our general approach for estimating treatment effects will be to regress each endline outcome on village treatment assignment, the baseline value of the outcome when available, baseline child age, baseline child gender, and randomization-strata fixed effects. The preferred specification is:

$$Y_{iv1} = \alpha + \beta T_v + \gamma Y_{iv0} + \delta' X_{iv0} + \lambda_s + \varepsilon_{iv}$$

where Y_{iv1} is the endline outcome for child-caregiver pair i in village v; T_v indicates assignment of village v to FIND; Y_{iv0} is the baseline value of the same outcome when available; X_{iv0} includes baseline child age and gender; and λ_s denotes county-level randomization-strata fixed effects. The coefficient β is the intention-to-treat (ITT) effect of FIND. For outcomes without a baseline analogue, Y_{iv0} will be omitted. Standard errors will be clustered at the village level. We will use Romano-Wolf stepdown adjusted p-values within each of the two primary behavioral outcome families and across the four intermediate-domain indices; we will not adjust across conceptually distinct families. Secondary and

exploratory outcomes will be clearly labeled and reported with unadjusted p-values, with family-adjusted p-values additionally reported when a secondary family contains multiple outcomes (Romano and Wolf, 2005).

Dealing with Missing Data and Attrition

We will address missing data in two ways. First, in cases of missing values for controls we will create a missing value dummy variable and include it in the regression to not lose that observation. We will also check that the results are not sensitive to this treatment of missing values comparing them to excluding those observations. Second, observations with missing data for the dependent variable will be discarded from the analysis (listwise deletions).

The primary specification will include only the prespecified baseline covariates. Robustness specifications may add an expanded set of pre-randomization child and household covariates and outcome-assessor fixed effects.

The primary ITT analysis will use all randomized observations with nonmissing endline outcomes for the outcome being analyzed; endline outcomes will not be imputed. We will compare overall and differential attrition across treatment arms and test whether treatment assignment predicts endline observation. As sensitivity analyses (Cunha et al., 2025), we will report inverse-probability-weighted estimates using a prespecified model based on treatment assignment, randomization strata, and baseline covariates. For outcomes with treatment-differential observation rates, we will also report Lee bounds when the monotonic-selection assumption is substantively plausible (Lee, 2009).

Intermediate Outcomes and Hypothesized Channels

We will first estimate the intent-to-treat effect of FIND on each of the four prespecified intermediate-outcome domains without conditioning on the other post-treatment domains. These estimates test whether FIND changes the caregiver-level outcomes predicted by the theory of change, they are not interpreted as causal mediation estimates.

The four domains are conceptualized as complementary intermediate outcomes: a). FIND's five core elements are expected to strengthen caregivers' beliefs about the returns to responsive interaction and their perceived capacity to engage responsively. b). Personalized, strengths-based video feedback may be particularly important for strengthening parenting self-efficacy and caregivers' relational perceptions of their children. c). Repeated practice may consolidate caregivers' perceived competence and support the translation of parenting beliefs into responsive behavior. d). Reductions in parenting stress may increase caregivers' psychological and attentional capacity to use these inputs during everyday caregiver-child interactions. We will interpret these patterns as evidence consistent with the hypothesized channels.

Heterogeneous Effects

We will explore heterogeneous treatment effects using the following prespecified baseline moderators: child age, child gender, an indicator for the mother being the primary caregiver, caregiver parenting beliefs, parenting self-efficacy, caregiver psychological stress/capacity, and caregiver-child relationship. Each moderator will be entered separately with its interaction with treatment assignment;

continuous moderators will be standardized using the full baseline sample. The interaction coefficient will be the estimand of interest. These analyses will be labeled exploratory, and additional subgroup analyses will be clearly distinguished as post hoc.

Treatment Take-Up and Implementation

We will report intervention attendance, number of completed home visits, treatment fidelity, participation, engagement, and available acceptability/barrier measures. These are process and implementation measures rather than psychological mediators. When noncompliance is meaningful, we present estimates of the dose–response relationship between the number of completed household visits and our main outcomes of interest, using randomized village assignment as an instrument for program participation.

References

Balldin, S., Fisher, P.A. and Wirtberg, I., 2018. Video feedback intervention with children: A systematic review. *Research on Social Work Practice*, *28*(6), pp.682-695.

Cunha, F., Gerdes, M., Hu, Q. and Nihtianova, S., 2024. Language environment and maternal expectations: An evaluation of the LENA start program. *Journal of Human Capital*, *18*(1), pp.105-139.

Cunha, F., Nihtianova, S., Rood, J. and van der Merwe, A.L., 2026. *Beliefs, Attention, and Investments in Early Childhood* (No. w35150). National Bureau of Economic Research.

Cunha, F., Hu, Q., Salvati, A., Wolpin, K.I. and Zeng, R., 2025. An evaluation of the Alief Independent School District Jumpstart program: using a model to recover mechanisms from an RCT (No. w33537). National Bureau of Economic Research.

Emmers, D., Jiang, Q., Xue, H., Zhang, Y., Zhang, Y., Zhao, Y., Liu, B., Dill, S.E., Qian, Y., Warrinnier, N. and Johnstone, H., 2021. Early childhood development and parental training interventions in rural China: a systematic review and meta-analysis. *BMJ Global Health*, *6*(8), p.e005578.

Lee, David S. 2009. "Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects." *Review of Economic Studies* 76(3): 1071–1102.

Liu, Sihong, Tiffany Phu, Amy Dominguez, Eliana Hurwich-Reiss, Drew McGee, Sarah Watamura, and Philip Fisher. "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* 26, no. Suppl 1 (2025): 52-65.

Romano, Joseph P., and Michael Wolf. 2005. "Stepwise Multiple Testing as Formalized Data Snooping." *Econometrica* 73(4): 1237–1282.

Table 1: Balance Check of Baseline Characteristics

Panel A.

VARIABLES	(1) Child age (months)	(2) Premature birth	(3) Male child	(4) Firstborn child	(5) CDI score (z)	(6) CREDI score (z)	(7) Mother is primary caregiver	(8) Caregiver age (years)	(9) Household asset index
Treatment (FIND)	0.402	-0.006	0.006	0.021	0.038	-0.001	0.004	-0.084	0.021
	(0.426)	(0.019)	(0.034)	(0.039)	(0.074)	(0.068)	(0.040)	(1.093)	(0.056)
Constant	21.489***	0.089***	0.566***	0.438***	-0.016	-0.003	0.517***	44.861***	-0.027
	(0.314)	(0.014)	(0.023)	(0.028)	(0.053)	(0.048)	(0.027)	(0.787)	(0.036)
Observations	827	824	827	827	818	822	827	814	815
R-squared	0.004	0.013	0.007	0.008	0.018	0.016	0.004	0.003	0.001

Panel B.

VARIABLES	(10) Caregiver ≥ 9 yrs schooling	(11) PSOC total (z)	(12) Read books with child	(13) Told stories to child	(14) Sang songs to child	(15) Played with child	(16) Number of books	(17) MORS warmth (z)	(18) MORS invasion (z)
Treatment (FIND)	-0.032	-0.067	-0.049	0.001	-0.016	-0.027	2.010	-0.010	0.054
	(0.035)	(0.069)	(0.042)	(0.037)	(0.032)	(0.027)	(1.217)	(0.075)	(0.070)
Constant	0.587***	0.034	0.524***	0.452***	0.748***	0.844***	13.877***	0.005	-0.032
	(0.026)	(0.052)	(0.031)	(0.028)	(0.021)	(0.021)	(0.809)	(0.055)	(0.044)
Observations	815	826	827	827	827	827	818	827	827
R-squared	0.007	0.019	0.008	0.003	0.002	0.008	0.027	0.006	0.006

*Notes: (a) Cluster-robust standard errors in parentheses; (b) *** p<0.01, ** p<0.05, * p<0.1; (c) a joint test of all baseline covariates shows no significant difference (at the 10% level) between the treatment and control arms.*