Violent Discipline and Child Behavior: The Short- and Medium-term Effects of Virtual Parenting Support to Caregivers in Jamaica*

Lelys Dinarte-Diaz† Saravana Ravindran‡
Manisha Shah§ Shawn Powers¶ Helen Baker-Henningham||

December 2022

Abstract

Despite the importance of quality parenting for child development, parenting practices are far from ideal, with 300 million 2-to-4-year-old children regularly subjected to violent discipline by caregivers worldwide. We experimentally study a virtually-delivered, scalable information intervention on positive parenting practices in Jamaica. The 10-week program was delivered through weekly SMS messages at a cost of only USD 1.43 per caregiver targeted. The SMS messages were complemented with access to an app with program materials and virtual information sessions with officers from Jamaica's Early Childhood Commission. In the short-term, the intervention succeeded in improving caregivers' knowledge (0.52 SD) and attitudes (0.2 SD) regarding parenting practices. This translated into changes in caregivers' disciplining behaviors, with a 0.12 SD reduction in violence against children. These results persist nine months after the end of the intervention. Furthermore, we find reductions in caregiver depression (0.12 SD), anxiety (0.16 SD), and parental stress (0.16 SD) in the medium-term. Treated children also faced fewer emotional problems (0.17 SD) in the short-term. The results can be explained by improvements in caregivers' self-efficacy (0.21 SD) and are not due to experimenter demand effects or displacement of violent disciplining to other children in the household. The results have scalable policy implications for human capital accumulation in developing countries.

Keywords: Mental health, positive parenting, parental stress, child maltreatment.

JEL Classification: J13, J22, I24, I12, J12, J167

*We are very thankful to Emanuela Galasso, Kathleen Beegle, and Victoria Baranov for their time and feedback. We also appreciate the comments from those who participated in seminars at the Development Research Group (World Bank), University of Exeter, Lee Kuan Yew School of Public Policy (NUS), Pontificia Universidad Catolica de Chile, Universidad Los Andes (Chile), and Universidad Andres Bello (Chile). We would also like to express our gratitude to Carlos Santiago Guzman, Gabriela Lecaro, Deyanara Matos, and Nathy Andrade for their superb work as research assistants. We are indebted to and truly grateful for the incredible support of the Early Childhood Commission (Jamaica) for implementing the intervention. This work was supported by the World Bank Early Learning Partnership Trust Fund and the Research Support Budget. Ravindran is funded by a startup grant at the Lee Kuan Yew School of Public Policy, National University of Singapore. This research project's protocol was reviewed and approved by the Institutional Review Board (IRB) at the University of West Indies in Jamaica in June 2021 with approval ID No. CREC-MN.86, 20-21 and by the Innovations for Poverty Action Research Ethics Committee in August 2022 with approval ID No. 15586. The trial was registered at the American Economic Association RCT registry - AEARCTR-0008266. The authors have no conflicts of interest to report. The findings, interpretations, and conclusions expressed in this report are entirely those of the authors. They do not necessarily represent the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent; nor those of the World Bank and its affiliated organizations, its Executive Directors, or the governments they represent.

†Development Research Group. The World Bank. Email: ldinartediaz@worldbank.org

‡Lee Kuan Yew School of Public Policy, National University of Singapore. Email: saravana@nus.edu.sg

§Luskin School of Public Affairs, UCLA, and NBER. Email: manishashah@ucla.edu

¶The World Bank. Email: spowers@worldbank.org

|| University of West Indies and Bangor University. Email: h.henningham@bangor.ac.uk

1 Introduction

Child abuse is strikingly common and remains a worldwide challenge. Around three in four children aged 2 to 4 worldwide—close to 300 million—are regularly subjected to violent physical punishment and/or psychological aggression by their caregivers at home (UNICEF, 2017).[1] Such exposure to violence can hinder children's development and undermine their sense of self-worth in the short term (Boden et al., 2007; Fry et al., 2018; Mersky and Topitzes, 2010). Moreover, research has shown that those who are victims of abuse and neglect as children are more likely to exhibit risky behaviors as teenagers (Hamby et al., 2011); to be absent from school, and have higher levels of aggression, mental distress, and social problems Lansford et al. (2002); and negatively affect their peers' test scores and behavior in the classroom (Carrell and Hoekstra, 2010). Lastly, early exposure to violence can have negative effects on outcomes during adulthood (Doyle Jr and Aizer, 2018), such as worse labor market outcomes (Currie and Spatz Widom, 2010) or involvement in crime (Currie and Tekin, 2012; Sviatschi, 2022). Overall, the high prevalence of child maltreatment and its potential long-term impact on children's well-being call for innovative and effective strategies to address this issue.

Evidence shows that parenting styles and decisions are crucial in driving human capital accumulation (Olivetti and Petrongolo, 2017; Doepke et al., 2019; Attanasio et al., 2020b).[2] Yet, parenting practices still promote violence against children, which is often rationalized as necessary or inevitable (UNICEF, 2017). In this sense, a potential intervention to reduce violence against children is parenting programs. Rigorous small in-person parenting programs oriented to improve children's development and promote positive interactions with their caregivers have shown promising results Jeong et al. (2019). However, the evidence about the effectiveness of violence prevention parenting programs at scale is scarce. Moreover, the effects of shifting the delivery of a parenting program from an in-person to a digital format are scarce and ambiguous.[3]

In this paper, we fill this gap of evidence by adapting and evaluating the short- and medium-term impacts of an early childhood, violence-prevention parenting program—the Irie Homes Toolbox (IHT)—for virtual delivery in Jamaica.[4] We measure the impacts of the intervention on caregivers' attitudes and be-

[1] This exposure to violence can even start earlier. According to data from 30 countries, nearly half of children aged 12 to 23 months are subjected to corporal punishment and a similar proportion are exposed to verbal abuse at home (UNICEF, 2017)

[2] Good parental inputs during early life stages influence the development of children's cognitive and socioemotional skills, which subsequently influence their health, school performance during adolescence, and labor market outcomes in adulthood (Carneiro et al., 2019; Attanasio et al., 2020b,a; Baranov et al., 2020).

[3] For example, Amaral et al. (2021) finds that a virtual program that gave caregivers tools and techniques to reduce negative discipline towards their children and to manage stress had unintended consequences on male caregivers stress.

[4] The in-person version of the program consists of an eight-week, group-based, parenting intervention, delivered through community preschools. Francis and Baker-Henningham (2020) shows that the program reduces parents' use of violence against their child, increases caregiver involvement with their child, and

haviors related to physical and psychological violence against their children. Moreover, we study the effect of the intervention on secondary outcomes, such as child behavior and emotional problems and caregivers' mental distress (stress, anxiety, and depression).

The virtual IHT is a 10-week program that consists of three SMS messages per week (for a total of 30 SMS messages). Each SMS message briefly describes the techniques caregivers were going to practice from each of the four core elements of the intervention: i) building positive relationships between parent and child, ii) preventing misbehavior, iii) managing misbehavior, and iv) supporting their children's emotional self-regulation (Francis and Baker-Henningham, 2020). The SMS messages cost only USD 1.43 per caregiver targeted, highlighting that our intervention was extremely cost-effective and scalable. These SMS messages were complemented with access to a data-free app with weekly content uploaded and weekly, one-hour, virtual group parenting sessions. These virtual group sessions were led by field officers of the Early Childhood Commission in Jamaica (ECC).

To measure the effects of the program, we collected self-reported information through a phone-based survey in three rounds: before the intervention, right after it was completed, and 9 months later. We also collected individual and household characteristics before the intervention started. Moreover, in each data collection round, we gathered data from caregivers (attitudes and perpetration of violence against children, caregiver's mental distress, parental self-efficacy, changes in their support networks, and caregiver involvement in play and learning activities), their children (individual characteristics and conduct and emotional problems), and other household members (individual characteristics).[5] Moreover, we collected administrative data on SMS messages reception and App usage from the phone company and reports on attendance at the virtual sessions provided by ECC field officers.

We document a strong first stage of our information intervention. 91% of caregivers in the treatment group reported receiving SMS messages as part of the intervention. 97% of these caregivers read the messages, and 98% of the caregivers who read the messages found them to be useful. Importantly, caregivers in the treatment group scored 0.53 SD higher ($p < 0.01$) on the information module relative to caregivers in the control group at the first follow-up. These treatment impacts on caregiver knowledge are persistent: the treatment group scored 0.39 SD higher ($p < 0.01$) on the information index relative to caregivers in the control group in the second follow-up.

The impacts on caregiver knowledge led to significant positive impacts on caregivers' attitudes to violence against children (VAC). Caregivers in the treatment group scored 0.2 SD lower on the attitudes to the VAC index ($p < 0.01$) in the short run. This is driven by a 0.19 SD decrease in the physical VAC sub-index

reduces child conduct problems among children with high levels of conduct problems at baseline.

[5] As we explain later in the paper, some modules were asked at the three rounds of data collection, whereas others were included in either the first or the second follow up surveys because we decided to prioritize the collection of other variables to measure mechanisms.

($p < 0.01$) and a 0.12 SD decrease in the psychological VAC sub-index ($p < 0.05$). These effects are persistent, with caregivers in the treatment group scoring 0.14 SD lower on the attitudes to the VAC index ($p < 0.05$) in the medium-run.

The change in attitudes to VAC among caregivers in the treatment group led to changes in their child disciplining behaviors. Caregivers in the treatment group scored 0.12 SD lower at the first follow-up on the violence against target child (VATC) index ($p < 0.05$). We construct sub-indices for the physical and psychological VATC and find 0.14 SD ($p < 0.05$) and 0.1 SD ($p < 0.1$) reductions in these indices in the short term, respectively. Furthermore, while we do not observe statistically significant impacts on caregivers' well-being at the first follow-up, we show intent-to-treat reductions of caregiver depression (0.12 SD, $p < 0.1$), anxiety (0.16 SD, $p < 0.05$), and parental stress (0.16 SD, $p < 0.05$) at the second follow-up. The changes in disciplining practices led to improvements in child outcomes. We find that reference children in the treatment group scored 0.17 SD lower on the index of the emotional problem in the short term ($p < 0.01$).

We address potential concerns relating to experimenter demand effects and displacement of violence toward other children in the household. First, we substantiate the unlikelihood that our results are due to experimenter demand. Following Asadullah et al. (2021) and Dhar et al. (2022), we test whether the treatment had an effect on the caregivers' social desirability index.[6] We do not find evidence that desirability bias changed among treated caregivers relative to those in the comparison group or that there were differential effects in our main outcomes between treated and control caregivers by having an above-median social desirability index. In additional robustness checks, we also include the social desirability index measure as an additional control variable. This test did not impact the coefficients and standard errors.

Second, we establish that the intervention did not displace violence toward other children in the household. Including data on the eldest child aged 7-12 in the household, we find 0.14 SD and 0.15 SD reductions in the index for violence against children in the short- and medium-term, respectively ($p < 0.01$). Third, we show that our study attrition is low and that although there is differential attrition between treatment and control groups in the first follow-up, most of our results are robust to correction for Lee bounds (Lee, 2009). Finally, we verify the robustness of our results to the inclusion of additional control variables as selected by a double LASSO algorithm. Our results remain similar in terms of magnitude and inference.

We explore two mechanisms for the impacts we find. We hypothesize that the intervention may have improved caregivers' (i) self-efficacy and (ii) support networks. shows positive impacts on caregivers' self-efficacy relating to self-acceptance at the second follow-up (0.21 SD, $p < 0.01$). We do not find statistically significant impacts on caregivers' parenting support or borrowing networks. The positive treatment effects and analysis of mechanisms suggest that the intervention provided caregivers with the necessary parenting

[6]This index captures the study participant's individual-level propensity to misreport sensitive items, which indicates whether or not the respondent is driven by the need for social approval.

tools and boosted their confidence in their own parenting skills. Taken together, our results suggest that the harsh behaviors of the caregivers could be explained, to a large extent, by a lack of information and preparedness.

Our paper contributes to several strands of the literature. First, our paper contributes to the evidence that measures the impact of light-touch and low-cost parenting programs on parenting skills and children's outcomes. Within this literature, few studies have assessed the impacts of providing parenting advice via text messages on children's outcomes in early childhood and caregivers' parenting practices. These studies present mixed results; some papers have shown that text message-based parenting programs can successfully improve parental engagement (Barrera et al., 2020) and preschoolers' cognitive and social development and learning (Widen et al., 2020; York et al., 2019; Hurwitz et al., 2015). Effects vary based on the content, personalization, and frequency of the text messages (Cortes et al., 2019; Doss et al., 2019). However, other studies have found either no effects of text messages-based parenting interventions on children's development or unintended consequences on caregivers' outcomes. For example, Barrera et al. (2020) do not detect effects on child development, and Amaral et al. (2021) finds that an SMS message-based parenting program that combines positive parenting and support to cope with mental distress can increase male caregivers' stress and reduce the interactions with their children. Our study adds to this literature by offering evidence of how a digital (information-based) program focused on non-violent parenting advice with very low implementation costs can reduce caregivers' violence toward their children and mental distress and children's emotional problems.

Second, we contribute to the evidence on child caregiving through parenting programs. A large body of evidence studies parenting programs designed to expose caregivers to parenting advice through intensive and direct means, such as home visits, with the objective of supporting children's health, nutritional, and cognitive deficiencies. A recent report on experimental evidence from early childhood stimulation programs in low- and middle-income countries confirms positive impacts on early childhood development (JPAL, 2020). Yet, the effect sizes from these studies vary substantially, and delivering effective interventions at scale has proved challenging (Jeong et al., 2018; JPAL, 2020). We add to that literature by testing an information-based non-violent parenting program, which shows positive effects not only on caregivers' violent attitudes and behaviors towards their children but also improvements in children's emotional problems.

Third, we also contribute to a small and nascent literature on digital and low-cost parenting programs targeted to reduce violence-based disciplining practices. In low-income settings, the widespread availability of mobile phones, the high incidence of violence against children, and the general acceptance of this issue make such tools attractive, but their effectiveness is not yet well understood. The two existing studies related to ours show ambiguous effects. First, Amaral et al. (2021) studies an information-based bundled intervention that, in addition to violence prevention practices, the program exposed caregivers to information

on stress management techniques. Similarly, Francis and Baker-Henningham (2020) studies the in-person version of the Irie Homes Toolbox. We show that this light-touch violence-prevention parenting program has important effects on caregivers' violence against their children that persists over time. Moreover, our study provides evidence of indirect positive effects of the intervention on caregivers' mental distress and parental self-efficacy.

Lastly, this paper complements a broader strand of literature that studies how parenting programs can be embedded in other social programs to help with their scalability and compliance. For example, cash transfer programs have led to shifts in parental investment and gains in cognitive development in several developing countries (Macours et al., 2012; Paxson and Schady, 2010; Levere et al., 2016; Attanasio et al., 2014). Other approaches used in the literature to address the scalability of the parenting program is to implement it in different settings, such as in health care services, (Mehrin et al., 2022; Hamadani et al., 2019), or by directly showing the feasibility of replicating the intervention at scale (Heckman et al., 2020). Our approach, however, is to study the question of scalability by testing an alternative and low-cost delivery mode of parenting intervention.

2 The Intervention

The Irie Homes Toolbox (IHT) consists of a violence prevention program targeting parents of 2-to-6-year-old children (hence correspondingly, from here on, we shall use the term "target child" to refer to the eldest child in a household in the 2-to-6-year age range). The content of the program is based on evidenced informed parenting practices that improve parenting behavior and reduce child behavior problems (Chorpita and Daleiden, 2009; Garland et al., 2008).[7] With the aim of scalability, the IHT was specifically designed to be integrated into the services provided by Jamaican preschools, suitable for an in-person implementation by preschool teachers, and to require few resources and equipment.

This face-to-face intervention includes content related to four key concepts: i) building positive relationships between parent and child (e.g. praise, child-led play, involving the child in everyday activities), ii) preventing misbehavior (e.g. understanding why children misbehave, giving children independence and autonomy, giving clear instructions, setting rules and expectations, modeling appropriate behavior), iii) managing misbehavior (e.g. redirecting child, withdrawing attention, setting limits, giving appropriate consequences) and iv) child emotional regulation and stress reduction techniques (Francis and Baker-Henningham, 2020).[8]

[7] For further details on the development of the IHT, see (Francis and Baker-Henningham, 2020).

[8] The face-to-face intervention also includes content related to supporting children's homework. This was originally included in the virtual intervention but the team removed it as we had to cut the duration of the intervention from 12 weeks to 10 weeks to improve its potential for scalability.

In this paper, we study a digitally adapted version of the Irie Homes Toolbox (vIHT). Studying the effectiveness of a digital parenting program is relevant for at least two reasons. First, after the pandemic began and subsequent stay-at-home orders were issued, traditional care services and caregiver support were not available. This fact was confirmed in Jamaica through a rapid-response survey conducted by our team. While the majority of caregivers in the survey sample reported increasing positive interactions with their children, they also reported more negative interactions. Second, the effectiveness of shifting the mode of delivery from an in-person to a digital one is not clear (Amaral et al., 2021; Linhares et al., 2022).[9] On the one hand, in-person group sessions may generate a sense of team spirit and peer-to-peer social support more easily than virtual sessions. However, digital interventions are easier to be implemented at scale and may allow participants to consume content and activities at their own pace. Lastly, the context of Jamaica is relevant for the virtual delivery of an intervention. Data from the Office of Utilities Regulation (https://our.org.jm/sectors/telecommunications/telecommunications-market-information-data/) indicates that the mobile subscription penetration rate was 104% in 2021.

The adaptation process of the IHT from an in-person to a virtual delivery consisted of four stages. The first stage involved adapting the materials and methods used in the face-to-face version to a virtual delivery format (SMS messages and filming videos); the second stage involved developing the resources for facilitators to implement the program including a full-scripted training manual and presentations; the third stage included activities related to piloting of individual sessions with two groups of 6 to 8 parents and making ongoing revisions based on lessons learned, and the fourth stage involved training specialists to lead the virtual groups.[10] During this process, the research team revised intervention content and materials based on participants' and specialists' suggestions and experiences documented throughout the process. Overall, this new digital format serves as an innovative alternative to complement other more costly parenting policy initiatives such as home visitations and group sessions.

The vIHT content was delivered over a 10-week period.[11] Beneficiaries received three SMS messages per week (30 SMS messages in total) relating to content from the Irie Homes Toolbox.[12] Each SMS briefly describes the techniques they were going to practice during the week. It also includes a link to the content embedded in the program App (see below) and information on the Irie Challenge for the week. The Irie Challenge consists of suggestions on how to put the strategies learned from the program into practice.

These SMS messages were complemented with two additional activities. First, beneficiaries also re-

[9]For example, Amaral et al. (2021) finds unintended effects of an information-based virtual parenting program on male caregivers' outcomes in El Salvador. Results show that post-intervention, participant male caregivers increased their stress and anxiety levels and reduced their positive interactions with their children.

[10]See Francis et al. (2022) for a description of the process followed for the adaptation of materials and video recording.

[11]See the content delivered during each weekly session in table A1 in the appendix.

[12]See the complete list of SMS messages in Table A2 in the Appendix

ceived access to a data-free App. This App included videos of parents utilizing the strategies with their children, the Irie Tower (which depicts the list of all strategies thought during the program), and the Irie Challenge. The content was uploaded and updated every week during the 10 weeks.[13] The second complementary activity consisted of a weekly virtual group with an early childhood education specialist. We partnered with the Early Childhood Commission (ECC) in Jamaica for the implementation of this third component. Caregivers were offered the opportunity to join a virtual parent group that met once per week for twelve weeks through GoogleMeet video-calls to discuss the specific topic of the week. These groups were between 8 to 9 participants and were formed randomly (see more details in the research design section).[14] During these sessions, the messages received by text and video were reinforced through discussion and practice. ECC officers were trained in the curriculum and its implementation by the Irie Toolbox Team based in the Caribbean Institute for Health Research.[15]

3 Experimental Design and Data

3.1 Participant Recruitment and Enrollment

We recruited participants through three channels: (i) sending SMS messages to Digicel's customers,[16] (ii) recruitment through the ECC and preschool principals, and (iii) recruitment through social media. We partnered with TrendMedia (Digicel) to send every participant a link to an enrollment survey through these three channels. For example, the link was included in the SMS messages sent to Digicel's customers, in a WhatsApp sent by preschool principals, or in the Social Media and Loop Campaign.[17] This survey included questions that helped us identify our target group. The survey included the following eligibility criteria: caregivers had to 1) live in the same house with at least one child 2 to 6 years old, 2) have access to a smartphone or tablet, and 3) provide consent to participate in the intervention and study. We also asked about gender and parish of residence in the enrollment survey, but these two variables were not eligibility criteria.[18]

[13] See figure A2 in the Appendix with examples of materials corresponding to Week 4.

[14] Participant caregivers received data packages that allowed them to join the virtual groups.

[15] Training involved eight hours of initial training split over three days, with groups of fifteen ECC officers followed by weekly training and support for two hours a week, in groups of 15 officers, for ten weeks. During these weekly training, the weekly sessions were demonstrated and practiced and the officers discussed opportunities and difficulties they were facing with caregivers during the implementation of the IHT training to get support from the team.

[16] As we show in Table A4, overall our main recruitment channel was through Digicel's customer base: 93.4% of participants were recruited via this channel.

[17] See SMS messages Snapshot in figure A3

[18] See snapshots of the full survey in figures A4 and A5.

As Figure 1 shows, a total of 2,396 eligible caregivers based on the criteria mentioned before. As we explain below, we collected baseline data from 1,113 individuals distributed across 14 parishes (46% of the total number of individuals enrolled and eligible).[19]

Figure A6 in the Appendix shows the final distribution of our sample compared to the distribution of the population across parishes. Overall, there seems to be a similarity in the proportion of participants in our sample and in the population in the parish between most of them except in Kingston and Saint Andrew (31.5% vs 24.6%), and Saint Catherine (29.6% vs 19.1%). We note that differences in enrolment rates can be explained by the geographic distribution of subscribers of the mobile operators. Overall, the SMS messages campaign was standard and did not target particular parishes.

3.2 Randomization

We randomly assigned the sample of 1,113 enrolled caregivers who met the eligibility criteria and completed the baseline survey to either the treatment or the control group with equal probability. That is, our sample includes 557 caregivers in the treatment group and 556 caregivers in the control group. Details on the intervention that the treatment group receives are provided in the intervention section. Caregivers in the control group receive three SMS messages per week (30 SMS messages in total) with content related to good practices to avoid COVID-19.[20]

Randomization was stratified by four strata, that is the cross between the gender of the caregiver (male or female) and the mode of recruitment into the study (SMS messages campaign or ECC/Principal referral and social media campaign). Appendix Table A4 indicates the size of each sampling stratum. Since only 7% of caregivers were recruited through the ECC/principal and social media channels, we pooled these two modes in the stratification exercise. Given the different roles the two sexes play in caregiving, we wanted treatment and control groups to be balanced in this regard. In addition, stratified randomization by gender will allow us to study heterogeneity in intervention impacts by the gender of the caregiver. Similarly, since the individuals who we contacted through the different enrollment channels could differ in characteristics that could be relevant to our study (e.g., currently participating or previously participated in preschool), we wanted to control for those potential differences.

[19] After performing the randomization, one caregiver in the treatment group was dropped because she was later identified as an officer from the ECC.

[20] See the complete list of SMS messages sent to the control group in Table A7 in the Appendix.

3.3 Data

3.3.1 Data collection stages

In this paper, we collected data at baseline (before the intervention) and conducted two follow up rounds (one right after the intervention ended and another one 9 months later). In the three rounds, we followed the same best practices on survey protocols. For example, we trained enumerators in the content and structure of the baseline instrument and protocol and contacted each participant up to 10 times. Although self-completed surveys are cheaper to implement, phone-based surveys are more effective to increase response rates (Amaral et al., 2022). For this reason, we administered phone-based surveys in all data collection rounds. We piloted the survey instrument before conducting the baseline (between May and June 2021). All data collected during the three rounds were self-reported. To reduce the risk of respondent fatigue, we limited the instrument length to approximately 35 minutes in each round. Participants also received a small monetary incentive (US$2.50) to complete each of the three surveys.

Baseline Data Collection: In August 2021, we contacted all the 2,396 caregivers who were eligible to participate in the intervention and consented to enroll in the study[21] and collected data from 1,113 of them at baseline.

The baseline survey includes modules on the caregiver's characteristics and outcomes, including employment status and well-being; attitudes and perpetration of violence against children; parental support activities and self-efficacy; social networks; and economic anxiety. We also include questions to measure the child's behavior and other socio-demographic characteristics. Finally, we add a module that collects information on household socioeconomic conditions. Section 3.3.2 describes the information that we collected in more detail. Following information protection protocols, we stored collected data on the survey firm's private server. Access to the data was restricted to project staff and researchers.

First Follow-Up Data Collection: The intervention was completed at the end of November 2021, and we collected the first round of follow up data in December 2021. We timed the first follow-up survey to test the short-term effects of the program and to minimize attrition. We contacted over the phone all 1,113 respondents who took the baseline survey and met the eligibility criteria. We collected follow-up data from 985 caregivers (an 88.5% response rate). The structure of the first follow up survey was similar to that of the baseline survey. In the former, however, we included a module on content assimilation to measure if caregivers self-report comprehending concepts taught in the vIHT.

[21] The remaining 1,283 (54%) enrolled individuals did not complete the baseline survey for several reasons, including they did not provide a correct phone number; we were unable to reach them after the maximum number of attempts determined in the ethics protocol; they changed their mind and decided not to participate in the study, among others.

Second Follow-Up Data Collection: To measure the medium-term effects of the vIHT in our outcomes of interest, we collected a second round of follow up data in August-September 2022. We contacted over the phone all 1,113 respondents who took the baseline survey and collected data from 728 caregivers (a 65% response rate). The structures of the first and second follow up surveys were similar in most of the modules, except that in the second follow up we excluded questions to measure caregivers' social networks and economic anxiety; and included alternative measures of parental self-efficacy and mental health distress, vignettes to measure parental discipline, and a module on social desirability bias. Please see Table A3 in the Appendix for a list of modules included in the surveys for each data collection round.

3.3.2 Survey Instruments and Outcomes

Based on our study's theory of change, we first test if the participants read the SMS messages, attended the intervention, and used the app and if they learned the concepts taught in the program. Second, we analyze the impact of the intervention on caregivers' attitudes and the perpetration of physical and psychological violence against children. Third, we also measure if the intervention had indirect impacts on caregivers' mental distress measures and on the child's conduct and emotional problems. As the intervention could directly impact caregivers who perpetrate violence, we could expect that their children's behavior would improve. Finally, we also collected data to measure potential mechanisms driving the effects on the main outcomes, such as parental self-efficacy and changes in their support networks. Note that some modules were asked during the three rounds of data collection, whereas others were included in either the first or the second follow-up surveys because we decided to prioritize the collection of other variables to measure mechanisms. Table A3 details the survey modules and corresponding stages of data collection. Appendix Section A2 presents a detailed description of all outcomes and the survey instruments we used to measure them.

A. SMS viewership, attendance, App use, and learning

We collected additional information to measure the take-up or use of each program component from different sources. First, we asked caregivers if they received the SMS messages and how relevant the content was for them. Second, TrendMedia shared individual-level information on whether caregivers logged in to the App and the time (in minutes) they were connected to the App. Finally, ECC officers collected attendance data for the virtual groups for each caregiver. We also collected information to measure if caregivers learned some concepts and practices that were taught in the program in the follow-up rounds. We included two statements for each of the four key concepts taught in the intervention.[22] For example, to measure

[22]Building positive relationships between parent and child, preventing misbehavior, managing misbehavior, and emotional self-regulation and stress reduction techniques. See the intervention Section for more details.

learning of the "building positive relationships between parent and child" concept, we asked caregivers to what extent they agreed with the following statements "Praising children helps them learn to behave well" and "It is important that parents take some time every day to play with their child doing what their child wants." The response options for each statement were on a 1- to 5-point Likert scale (1–Strongly disagree, and 5–Strongly agree). We estimate a learning index—the greater the index, the more caregivers knew about the program content.

B. Main outcomes

Attitudes towards violence against children: We use an adapted version of the UNICEF MICS questionnaire to measure parental attitudes towards physical and psychological violence against children at baseline and follow-up rounds. The adapted instrument includes 5 items asking about some attitudes such as if they agree that a good parent can slap the child if he misbehaves, and if shouting and yelling would make the child more obedient, among others. Since the survey was over the phone and our goal was to conduct a 35-minute length survey, we only selected the 5 items (out of 13 items) with the greatest variation according to the results from the instrument piloting that was conducted before baseline data collection.[23] The greater the index, the more pro-violence the caregivers' attitudes.

Violence against children (self-reported): In the three rounds of data collection, we used a shortened version of the UNICEF MICS questionnaire to measure caregivers' perpetration of physical or psychological violence against children. The adapted instrument includes 5 items asking about some violent behaviors. These can be grouped into physical violence (hitting the child with a bare hand or with an object) or psychological violence (shouting, yelling, or screaming at the child; saying to send the child away; threatening to hit the child). We asked each caregiver about perpetrating these violent acts separately to the "target child" (eldest child 2 to 6 years old) or to another older child within the household (eldest child between 7 to 12 years old). Using these reports, we created two indexes: violence against the target child and violence against any child within the household. The latter is a pooled measure for both the target child and any other older child within the household. Moreover, using the items for each of the types of violence against the target child, we also created separate indexes for physical and psychological violence. The greater any of the indexes, the more violent acts were perpetrated by the caregiver against the target child or another child in the household.

C. Secondary Outcomes

Caregiver's mental health: We collected data to measure caregiver's depression using the Patient Health Questionnaire (PHQ-2 survey, Kroenke et al. (2003)) and a question on having difficulty sleeping at night. We also measure anxiety using the Generalized Anxiety Disorder (GAD-2, Donker et al. (2011)) instrument, re-

[23] Results from the pilot of the instrument showed a good test-retest and internal consistency of the 5 items selected to measure attitudes towards violence against children.

spectively. The PHQ-2 and GAD-2 questionnaires include two items each asking how often the caregiver had been bothered by any of the problems over the last two weeks. We included these instruments during the three rounds of data collection. Moreover, during the second follow-up, we also included the questions from the 18-items Parental Stress Scale (PSS-18, Berry and Jones (1995)). Our main outcomes of interest are the aggregate indexes of depression, anxiety, and stress separately. The greater the index, the higher the levels of depression or anxiety.

Child conduct and emotional problems: We use the 10 items related to children's conduct and emotional problems (5 items each) from the Strengths and Difficulties Questionnaire (SDQ) instrument to measure a child's behavior index. We collected this information in each of the survey rounds. Each question is answered on a 0-2 scale (Not true, somewhat true, certainly true). The items ask if some behaviors, which serve as proxies for conduct and emotional problems, occurred for the child during the previous 3 months. The greater the index, the more conduct or emotional problems the child has.

D. Mechanisms

Parental self-efficacy: We measure parental self-efficacy at baseline and during the first follow-up round using the 5 items from the Brief Parental Self Efficacy Scale (BPSES) instrument. The scale asks parents about their agreement with statements that can describe their ability to improve a child's behavior. For the second follow-up, we adapted and used the Tool to Measure Parenting Self-Efficacy (TOPSE) for more detailed questions relating to discipline and self-acceptance. In each survey wave, the main outcome of interest is the aggregate index of parental self-efficacy. The greater the index, the higher the self-assessment of efficacy.

Caregiver's support networks: As existing evidence shows, the effectiveness of positive parenting programs can be driven by the creation of support networks for participant caregivers. To test this potential mechanism in our context, we collected information on whether caregivers obtained support from friends, family, or professionals to solve parenting or financial issues. We asked how many people they could reach out to in case they need to talk about issues related to parenting and child rearing or borrowing money. Using this information, we created two indexes: one for parenting support and another for financial support. The greater the index, the larger the network the caregivers have.

E. Sociodemographic characteristics and other controls

Socio-economic and demographics. We collected the following socio-demographic data on the main caregivers: age, gender, education, marital status, employment status and occupation, income, household composition, and recent changes in lifestyle as a result of COVID-19 experiences. We also collected the age and gender of the target child. We collected information from all children aged 17 and below and their caregivers who regularly live under the same roof in the household. For both children and caregivers, we asked about their age and gender. In addition, we also asked about the education, marital status, employment, and occupation of all household members for each caregiver living in the household. All this information was provided by the caregiver enrolled in the study.

Social Desirability Bias: A potential concern with using self-reported data to measure sensitive outcomes such as violence is the experimenter demand effect. To account for this potential bias in our estimations, we included the short form of the Marlowe-Crowne Social Desirability Scale (Crowne and Marlowe, 1960). It consists of a survey module developed by social psychologists to measure a person's propensity to give socially desirable answers. The module, which we included in the second follow-up survey, asks respondents if they have several too-good-to-be-true traits such as never being jealous of another person's good fortune and always being a good listener. Using this information, we created an index of social desirability bias for each participant. The greater the index, the more socially desirable the responses of the participant.

Caregiver involvement in play and learning activities: We collect data to measure this outcome as a placebo for our analysis. As we explain later, this outcome helps us to address the concern of the experimenter demand effect. We use 5 items from the UNICEF's MICS questionnaire related to parental support for learning and play activities. The adapted instrument includes 5 items asking the caregiver if, during the past week, she played with the child if told stories, or sang songs to the child, and others. The items ask how often different activities occurred between the caregiver and the child during the previous week. This module was included in each of the data collection rounds. The index is estimated using the 5 items and was piloted before conducting baseline and showed a good test-retest and internal consistency. The greater the index, the more supportive the caregiver of her child's learning and play.

3.4 Baseline Summary Statistics

Table A5 reports the summary statistics of the variables used for the baseline. In panels A and B we present caregivers' and children's characteristics. Panels C - E show statistics of main outcomes (panel C), secondary outcomes (panel D), and mechanisms measures (panel E).

On average, caregivers are 33 years old, have 14 years of education, and 85% of them are female. Furthermore, 37% of the caregivers are married, 79% reported being employed and the average income in the past month was $137,308 JMD (USD 882 approximately). With respect to household characteristics, the average household has 4.6 members, with approximately 2 of them being children under 17, and approximately 24% of caregivers face food insecurity issues. Target children are, on average, 4.1 years old and are gender-balanced (49% are female).

On the other hand, caregivers' outcomes suggest harsh behaviors toward their children. The average caregiver draws on harsh conduct to discipline their children 2 to 6 years old approximately 1.3 times per week, and 1.1 times per week for the oldest children. Furthermore, caregivers exhibit relatively high support for violence against children, with 18.3% in agreement with violent attitudes. For child behavior-related outcomes (secondary outcomes), we observe that 41.8% of children exhibit conduct problems, while 25.9% display emotional problems. Moreover, the average caregiver in our sample has low symptoms of

depression and anxiety, with average prevalence rates of 19% and 15% respectively.

In terms of measures of mechanisms, on average caregivers could ask for help from approximately 2.5 relatives (among friends, family, and professionals). For parenting issues, the network is slightly bigger, with an average of 3 people, and with 16% of them participating in parent support groups. Parents, on average, interacted with their children by playing, reading books, telling stories, drawing together, or praising them 4.8 days per week.

How do caregivers in our sample compare to the representative caregiver of a child aged 2-6 in Jamaica? We compare our sample of 1,113 caregivers with the Jamaica Survey of Living Conditions (JSLC) 2019. The Jamaica Survey of Living Conditions is a living standards measurement survey that is representative of the Jamaican population. We restrict the JSLC sample to only caregivers with a child aged 2-6 and compare key demographic variables for which data is available across surveys in Appendix Table A6. The last column of the table provides p-values for the comparison of means across the two samples. We find that caregivers in our sample are younger, less likely to be female, and have one more year of education. They are also more likely to be married, be employed, and have fewer children. The share of female target children is similar across the two groups. Overall, caregivers in our sample seem to have better socioeconomic status than the average caregiver in Jamaica. Therefore, our research design suggests that one should be careful regarding the external validity of our results.

4 Estimation Framework

4.1 Validity of the experimental design

We use a randomized control trial (RCT) to identify the effect of IRIE Homes on our primary outcomes. The main identification assumption is the following: absent the intervention, our outcomes of interest would be, on average, statistically indistinguishable between caregivers assigned to the treatment and control groups.

To support the internal validity of the experiment, we conduct two analyses. First, we test for balance on pre-intervention variables between treatment and control groups. These results are presented in Table 1. Columns 1 and 2 presents the sample with valid information for each variable and the mean and standard deviation (SD) for treatment (control), respectively. Column 5 provides p-values for t-tests for equality of means between the treatment and control groups. With the exception of one out of 32 variables tested, we do not find significant differences in these variables across treatment and control groups at p-values less than 0.1. The only difference we find is that caregivers in the treatment group were more likely to suffer from high anxiety at baseline (12.9% in the treatment group versus 16.6% in the control group; $p = 0.084$). Statistically, however, this is not unexpected when testing across a large number of variables (31).

Reassuringly, the p-value for the overall F-test of joint orthogonality is 0.947, highlighting that jointly, the means of the variables are not statistically distinguishable across the treatment and control groups. Taken together, the randomization produced comparable treatment and control groups.

Second, we assess differential attrition between treatment and control groups for each follow up round and present the results of this analysis in Table A10. Column (1) shows that caregivers in the treatment group were 4.7 percentage points more likely to complete the first follow up survey relative to caregivers in the control group.[24] In column (2), we study whether the differential attrition in the first follow up was correlated with any demographic characteristics or outcome variables measured at baseline. None of the interaction terms in this regression are statistically significant at the 10% level, suggesting no evidence of relationships between demographic characteristics or outcome variables and the differential attrition. In line with this result, we cannot reject the null hypothesis that the all interaction terms between treatment indicator and the relevant variables are not statistically significant (p-value of 0.528).

In the second follow up, we find no differences in the probability of not completing the survey between treatment and control groups (column 3). Moreover, we find that none of the relevant variables explain any differential attrition between treatment and control, we do not reject the null hypothesis that the interaction terms between treatment indicator and all the relevant variables are not different from zero (p-value = 0.798).

4.2 Empirical methods

To study intent-to-treat (ITT) impacts, we estimate Ordinary Least Squares (OLS) regressions using the following Analysis of Covariance (ANCOVA) specification for caregiver i in period t and stratum s:

$$Y_{i,t} = \beta_0 + \beta_1 T_i + \beta_2 Y_{i,t-1} + \gamma_s + \varepsilon_{i,t} \tag{1}$$

where $Y_{i,t}$ refers to the outcome variable of interest of caregiver i as measured at first or second follow-up, defined in Section 3.3.2. T_i is an indicator variable capturing assignment of caregiver i to the treatment group, and $Y_{i,t-1}$ refers to the outcome variable of interest measured at baseline. γ_s captures stratum fixed effects for the four strata; the interaction between the gender of the caregiver (female or male) and the two modes of recruitment (SMS messages or social media). The estimate of β_1 captures the ITT impact of the treatment. We estimate and report heteroskedasticity-robust standard errors.

We conducted Least Absolute Shrinkage and Selection Operator (LASSO) analysis to identify variables with strong relationships with $Y_{i,t}$, to assess their suitability for inclusion as controls in Equation (1). This follows the recommendations of Bruhn and McKenzie (2009) that authors "should choose which baseline

[24] As a robustness check, present our key results adjusting for Lee (2009) bounds in Section 6.

characteristics to control for, not on the basis of statistical differences, but on the strength of their relationship to the outcome of interest." Since LASSO consistently only selected the outcome variable of interest measured at baseline ($Y_{i,t-1}$) for inclusion across all outcomes, we do not include other control variables in Equation (1).[25]

We performed power calculations for our key outcome variable on violence against children as part of our pre-analysis plan. Appendix A1 details these calculations and Figure A1 presents a graph of power under different assumptions. Under our most conservative assumptions, we are powered to detect ITT impacts of 0.18 SD, although this minimum detectable effect (MDE) could be smaller depending on the rate of attrition and correlation between baseline and follow-up.

To address potential concerns relating to multiple hypothesis testing, we construct indices for broad families of outcomes using Anderson (2008). Summary index tests offer three advantages: (i) they are robust to over-testing because each index represents a single test, (ii) they provide a statistical test for whether a program has a "general effect" on a set of outcomes, and (iii) they are potentially more powerful than individual-level tests by reducing random error in each outcome measure (Anderson, 2008). Each summary index is a weighted mean of several standardized outcomes, where the weights are calculated to maximize the amount of information captured in the index using an efficient generalized least squares (GLS) estimator.

5 Short- and Medium-term Results

Our theory of change hypothesizes that caregivers assigned to the treatment group received information on parenting practices relating to the vIHT, attended the weekly virtual sessions with ECC officers, and updated their beliefs on attitudes to violence against children (VAC). We hypothesize that they consequently changed their behaviors relating to VAC and parenting practices, and that caregiver well-being and child outcomes relating to conduct and emotional problems improved as a result. We test these hypotheses starting with first stage impacts on attendance and learning in Section 5.1, followed by impacts on caregiver attitudes and behaviors in Sections 5.2 and 5.3, respectively. We study caregiver well-being in Section 5.4 and child outcomes of interest in Section 5.5.

[25] Yet, LASSO does select other controls, which vary across outcomes. For example, for some outcomes it selects age and gender as additional controls, whereas for others, LASSO selects household composition and educational level of the caregiver. To confirm that our results do not change by the inclusion of these additional controls, we include these estimations as a robustness check in section 6.

5.1 First Stage: Program Take-up & Learning

As we highlight in Section 2, the intervention consisted of three components: three SMS messages per week, access to a data-free App with vIHT content, and weekly virtual sessions with ECC officers.

To understand receipt of SMS messages, App usage, and attendance at sessions, we present descriptive statistics on these items in Table 2. Panel A highlights that 91% of caregivers in the treatment group reported receiving SMS messages as part of the intervention. Of the 499 caregivers who reported receiving the SMS messages, 97% read them. Furthermore, 98% of the 444 caregivers who read the messages found them to be useful. Using streaming data from the App, we also tracked the total duration of time spent accessing content on the App in Panel B. The mean duration spent on the App (across the 10-week intervention) was 6.9 minutes across all treatment groups. Moreover, the mean number of sessions accessed on the App (out of 10) was 1.0. In Panel C, we show that across all caregivers assigned to the treatment group, the mean number of sessions attended (out of 10) was 4.6. In sum, although this intervention included two complementary components to the SMS messages, these take-up rates lead us to conclude that all the estimated impacts are driven mainly by the SMS messages rather than by any of the other two complementary elements.

To assess whether caregivers increased their knowledge relating to the vIHT content, we administered an information module to all caregivers at endline. Table 3 presents ITT impacts of the treatment on caregiver's knowledge. Columns 1 to 8 present the treatment impacts on eight statements relating to parenting practices. The eight statements were designed to assess understanding of the four key concepts of the intervention as outlined in Section 2 (two statements were asked for each of the four concepts). All statements are true; thus higher values represent greater knowledge. Panel A presents the short-term results from the first follow-up conducted immediately after the intervention ended. We see that caregivers in the treatment group were significantly more likely to state that the statements were true relative to caregivers in the control group for six out of the eight statements (four statements significant at the 1% level, one statement significant at the 5% level, and one significant at the 10% level). To address potential concerns relating to multiple hypothesis testing, we aggregated the responses to the eight statements into an information index as outlined in Section 4. Column 9 of Table 3 presents the ITT impacts of the treatment on the information index. Caregivers in the treatment group scored 0.53 SD higher on the information module relative to caregivers in the control group. The impact is large in magnitude and statistically significant at the 1% level. Panel B presents the medium-term results from the second follow-up conducted nine months after the intervention ended. We see persistent treatment impacts on caregiver knowledge: the treatment group scored 0.39 SD higher on the information index relative to caregivers in the control group ($p < 0.01$). Panel D of Table A14 shows that the medium-term impact on the information index is not statistically different from the short-term impact.

Does attendance at a greater number of sessions lead to more learning? Acknowledging that selection

into attendance at sessions is endogenous, we explore this dose-response relationship in Figure A7, where we present coefficient plots of the impact of session attendance on the caregiver information index. To improve precision, we group attendance into 5 pairs of possible combinations, i.e. attendance at one or two sessions, three or four sessions, and so on (where attendance at none of the sessions represents the base reference group). The regression is estimated over caregivers in the treatment group only. We see an increasing dose-response relationship: while caregivers attending one to four sessions did not score significantly higher than caregivers who did not attend any sessions, caregivers attending five to ten sessions scored significantly higher on the information index in the short-term relative to caregivers who did not attend any sessions. With the exception of attendance at seven or eight sessions, the caregiver's knowledge monotonically increases in the short term in the number of virtual sessions attended. Results from the second follow-up show a very similar pattern in the medium term, with caregivers attending seven to ten sessions scoring significantly higher on the information index relative to caregivers who did not attend any sessions.

Overall, we find that while attendance at the weekly virtual sessions was imperfect and App usage low, the intervention succeeded in increasing the caregiver's knowledge relating to the vIHT content and parenting practices.

5.2 ITT Impacts on Caregiver Attitudes to VAC

We present intent-to-treat estimates of the treatment impact on caregiver attitudes to violence against children (VAC) in Figure 2 (see Section 3.3.2 for more information on the indices; the greater the index, the more pro-violence are the caregivers' attitudes).

Figure 2 shows that caregivers in the treatment group scored 0.2 SD lower on the attitudes to (VAC) index ($p < 0.01$) in the short-run. Panel A of Table A8 presents results from the first follow-up and breaks down the attitudes to VAC into five individual components. The first three columns constitute the attitudes towards physical VAC, while columns 4 and 5 comprise the attitudes towards psychological VAC. In the short term, caregivers in the treatment group were 3.9 percentage points less likely to agree with the statement that children need to be physically punished in order to bring up, raise, or educate a child properly. The ITT impact is large relative to the control group mean of 12.1% who agreed with the statement (i.e., a 3.9/12.1 = 32% reduction) ($p < 0.05$). They were also 6.2 percentage points (27%) less likely to agree that a good parent slaps their child when they misbehave and 5.1 percentage points (53%) less likely to agree that when a child is beaten, he/she will stop doing the unwanted behavior ($p < 0.01$). The sub-index for attitudes to physical VAC is shown in Figure 2. In the short term, we find a 0.19 SD decrease in this sub-index among caregivers in the treatment group ($p < 0.01$).

In terms of short-run changes in attitudes towards psychological VAC, caregivers in the treatment group

were 2.3 percentage points less likely to agree that shouting and yelling makes the child more obedient and 2.7 percentage points more likely to disagree with the statement that "shouting, yelling, and threatening to slap will harm the child," although these results were not statistically significant at conventional levels. Figure 2 also presents the treatment impact on the sub-index we construct for attitudes to psychological VAC. At the first follow-up, we find a 0.12 SD decrease in this sub-index among caregivers in the treatment group ($p < 0.05$).

These changes in attitudes persist into the medium-term. Figure 2 shows that caregivers in the treatment group scored 0.14 SD lower on the attitudes to violence against children (VAC) index ($p < 0.05$) in the medium-run. Analyzing the sub-indices for attitudes to physical and psychological VAC, we see persistent impacts on attitudes to physical VAC (0.15 SD reduction, $p < 0.05$), although the changes in attitudes to psychological VAC are no longer statistically significant. Panels A and B of Table A14 shows that the medium-term impacts are not statistically different from the short-term impacts.

5.3 ITT Impacts on Caregiver Behaviors Relating to VATC

The change in attitudes to VAC among caregivers in the treatment group led to changes in their disciplining behaviors of their target child. The intent-to-treat estimates shown in Figure 3 highlight that caregivers in the treatment group scored 0.12 SD lower at first follow-up on the violence against target child (VATC) index ($p < 0.05$). The figure also shows the sub-indices that we construct for physical and psychological VATC. We find 0.14 SD ($p < 0.05$) and 0.1 SD ($p < 0.1$) reductions in these indices in the short term, respectively. The larger treatment impact on physical versus psychological VATC is in line with the larger treatment impact found on attitudes to physical versus psychological VAC. These results persist nine months after the end of the intervention: caregivers in the treatment group scored 0.13 SD lower at the second follow-up on the violence against target child (VATC) index ($p < 0.05$). In line with the results on attitudes to physical and psychological VAC in the medium-term, we find statistically significant reductions in physical VATC (0.12 SD reduction, $p < 0.1$), but do not find statistically significant impacts on psychological VATC. Panels A and B of Table A14 shows that the medium-term impacts are not statistically different from the short-term impacts.

Panel B of Table A8 presents the extensive margin short-term impacts for the five components that comprise the VATC index. The first two columns are components of the physical VATC index, while columns 3, 4, and 5 comprise components of the psychological VATC index. We find that caregivers in the treatment group were 9.7 percentage points (25%) less likely to hit their child on the bottom, hand, arm, or leg with their bare hand ($p < 0.01$). In terms of psychological VATC, caregivers in the treatment group were 10.1 percentage points (14%) less likely to shout, yell, or scream at their child ($p < 0.01$) and 5.5 percentage points (8%) less likely to threaten to hit their child without doing so ($p < 0.1$). However, they were 4 percentage

points (44%) *more likely* to say to their child that they will send them away (i.e. a time-out) ($p < 0.05$). This indicates that caregivers in the treatment group substituted away from other forms of discipline towards psychological discipline in the form of a threat of a time-out. Taken together, however, there is a net decrease in the VATC index.

Overall, we find that the treatment led to positive and persistent impacts on caregivers' attitudes towards VAC, which in turn led to persistent reductions in VATC. In the next section, we turn to the impacts this had on child behaviors and skills.

5.4 ITT Impacts on Caregiver Depression, Anxiety, & Stress

Harsh behaviors towards children may be explained by stress, anxiety, and frustration (Persson and Rossin-Slater, 2018; Bendini and Dinarte, 2020), which affect parental functioning through psychological well-being (Belsky, 1984; Belsky and Jaffee, 2006; Taraban and Shaw, 2018). We study intent-to-treat impacts on caregivers' well-being as measured by depression, anxiety, and parental stress. Depression and anxiety were measured at both follow-ups using the Patient Health Questionnaire (PHQ-2) and Generalized Anxiety Disorder (GAD-2) instruments, respectively. Parental stress was measured using the Parental Stress Scale (PSS-18) developed by Berry and Jones (1995) at the second follow-up only.

While we do not observe statistically significant impacts on caregivers' well-being at the first follow-up, Figure 4 shows intent-to-treat reductions of caregiver depression (0.12 SD, $p < 0.1$), anxiety (0.16 SD, $p < 0.05$), and parental stress (0.16 SD, $p < 0.05$) at the second follow-up. The lack of initial impacts suggests that only after caregivers were aware of and applied the new tools learned from the intervention, then their mental distress improved. Alternatively, caregivers may have needed time to apply their learning from the intervention. Section 7 discusses potential mechanisms in more detail.

5.5 ITT Impacts on Child Behaviors

Our theory of change hypothesized that positive impacts on caregiver attitudes and behaviors will lead to positive impacts on child behaviors. Figure 5 presents the intent-to-treat impacts of the intervention on the target child's conduct and emotional problems. As outlined in Section 3.3.2, conduct and emotional problems were measured using caregivers' responses to the Strengths and Difficulties Questionnaire (SDQ) instrument.

The estimates show that children in the treatment group scored 0.17 SD lower on the index of the emotional problem in the short term ($p < 0.01$). These children also scored 0.03 SD lower on the conduct problems index in the short term; however, this is not statistically significant at conventional levels. In the medium term, however, we do not find statistically significant reductions in conduct and emotional prob-

lems. Panel C of Table A14 shows that the medium-term impacts are not statistically different from the short-term impacts.

6 Addressing Alternative Interpretations of the Results

In this section, we address potential concerns regarding the results presented in Section 5, such as experimenter demand effects, displacement of violence toward other children in the household, and differential attrition across treatment and control groups in the first follow-up. We also present the results of additional robustness checks.

6.1 Assessing Potential Bias Due To Experimenter Demand Effects

Self-reported measures are susceptible to potential experimenter demand effects when assessing sensitive information such as attitudes and perpetration of violence because participants' responses regarding sensitive topics might be influenced by social desirability bias (Aguero and Frisancho, 2021; Amaral et al., 2021). The concern is whether participants report statements on violent attitudes and behaviors differently from their attitudes and behaviors outside of the study environment and whether this is differential across the treatment and control groups. This is unlikely because the SMS messages and other components of the intervention did not include content or words related to physical or psychological violence (see Table A2). As highlighted in Section 2, the focus was on providing techniques to help caregivers to build positive relationships with their children, prevent and manage children's misbehavior, and to help their children to regulate their emotions.

Yet, we take seriously the possibility of experimenter demand effects and conduct several exercises to address this potential concern. First, we test the intervention's direct effect on a social desirability index (SDI) estimated using the Marlowe-Crowne Social Desirability Scale (Crowne and Marlowe, 1960).[26] As we show in Column (1) Panel B of Table 4, there are no statistically significant differences between treated and control caregivers on the SDI. Second, we test the robustness of all of our results to the inclusion of the SDI as an additional control variable. These results are shown in Columns (2)-(5) of Table 4. The estimated effects remain similar in magnitude and statistical significance.

Third, we follow Dhar et al. (2022) to study whether potential experimenter demand effects may bias the estimated program impacts upward. A potential concern might be that caregivers in the treatment group have a higher propensity to give socially desirable answers, thereby biasing the estimated treatment effects.

[26]The social desirability measures were only collected during the second follow-up.

Our results in Table 5 show that there are no heterogeneous treatment impacts by SDI; the interaction terms in the third row are statistically insignificant.

Fourth, we study placebo outcomes – outcomes that the intervention was not set up to impact – as further evidence against experimenter demand effects. In particular, we study if participants in the treatment group were more inclined to report greater caregiver-child play and learning interactions. To study this, we conduct a placebo test and use as our outcome an index of caregiver involvement in play and learning activities with the child. To construct the index, we asked whether caregivers played, read books, sang songs, drew with, or praised their children in the past seven days. As we show in Table A9, the estimated effects for the overall index and each of its sub-components are not statistically different from zero. This result confirms that the caregivers in the treatment group did not simply tend to report improvement in any parenting-related outcome but only in those that were specifically targeted through the program.

Together, these findings suggest that it is unlikely that experimenter demand effects might have changed respondents' likelihood to systematically misreport their attitudes and behaviors about the use of violence with their children across the treatment and control groups.

6.2 Assessing Potential Displacement of Violence Toward Other Children

A potential concern might be that the intervention succeeded in reducing caregiver violent discipline against the target child, but this may have displaced violence toward other children in the household. To address this concern, we surveyed caregivers about violence against their eldest child aged 7-12 in addition to the target child. Panel A of Table A5 shows that households in our sample have, on average, 1.9 children aged 17 or below; thus, given the constraints of the phone survey, we collected information on one elder child in the household. We stack the responses on caregiver attitudes and behaviors to violence against children to form a child-level dataset that allows us to study violence against the target child, as well as the eldest child aged 7-12.

Figure 6 shows 0.14 SD and 0.15 SD intent-to-treat reductions in the index for violence against these children in the short- and medium-term, respectively ($p < 0.01$). We also find 0.1 SD ($p < 0.1$) and 0.13 SD ($p < 0.05$) reductions in the indices of physical violence in the short- and medium-term, respectively. Moreover, we also find that caregivers in the treatment group scored 0.13 SD ($p < 0.01$) and 0.14 SD ($p < 0.05$) lower in the indices of psychological violence against any of these children at the first and second follow-ups. Panels A and B of Table A14 show that the medium-term impacts are not statistically different from the short-term impacts.

Panel C of Table A8 shows that the results for each of the items included in the estimation of the indices are similar in magnitude and statistical significance to the ITT estimates in Panel B, which only focuses

on the target child, except for threatening to send their child away, which is smaller and statistically non-significant.

Overall, these results suggest that the intervention did not displace violence toward other children in the household.

6.3 Assessing Potential Bias Due To Differential Attrition

Another potential concern from our estimations is that the differential attrition between treatment and control groups during the first follow-up round may bias our estimates. In particular, if the higher proportion of treatment group caregivers who attrited were more likely to exhibit violent disciplining behaviors, our estimates may overestimate the true treatment effects. To address this concern, we estimate Lee bounds to account for sample selection (Lee, 2009) and present these results in Table A11. This procedure is a conservative estimate of the treatment effect, as it corresponds to extreme assumptions about the missing information. We find that all upper and lower bounds significantly differ from zero except the upper bound of the index of violence against the target child (p-value = 0.177), suggesting that our results are robust to differential attrition.

As an additional check, we re-estimated the ITT impacts on caregiver attitudes and behaviors, as well as child outcomes using a balanced panel of caregivers that were present at both the first and second follow-ups. This addresses concerns of potential selection in who might be present at either follow-up round. The results are shown in Figures A8 - A11. Overall, the results are very similar to the estimates presented in Section 5.

6.4 Assessing Sensitivity From Selection of Control Variables

As we discuss in Section 4, we use a double LASSO approach to identify the variables that should be included in our estimations as controls. We find that LASSO consistently selects, across all outcomes, the measure of the outcome at baseline. LASSO also select other variables that can be used as controls, but they vary across outcomes. For instance, for some outcomes age and gender are selected, whereas for others LASSO selects education level and household composition. Considering this, we test for the stability of our estimated coefficients after including the control variables selected by LASSO for each of our main outcomes. As we show in Table A12, the estimated coefficients and their statistical significance do not change after including these additional control variables selected by LASSO.

7 Mechanisms

Prior research has found that the lack of information or parenting preparation is a trigger of the previous cognitive burdens (Baker-Henningham and Francis, 2018), leading to harsh and unhealthy parenting behaviors. Indeed, responsiveness-oriented parenting approaches[27] have shown the importance of self-efficacy beliefs in parenting functioning strategies (Michl-Petzing et al., 2019). As more warmth and comprehensive parenting responses are fundamental for children's development (Landry et al., 2001, 2012), parents' perceptions about their self-efficacy (i.e. the extent to which they consider themselves capable and prepared to raise a child and deal with the associated parenting tasks) becomes a cornerstone in parenting good practices (Izzo et al., 2000; Hoover-Dempsey et al., 2005; Jones and Prinz, 2005). Self-efficacy, then, could be enhanced by mastering new experiences and by improving the provision of information related to parenting, reducing stress, and increasing the quality of the interactions between caregivers and the children (Bandura, 1982; Bandura and Wessels, 1994; Bloomfield and Kendall, 2012).

In this section, we explore two hypotheses for the impacts we find: we hypothesize that the intervention may have improved caregivers' (i) self-efficacy and (ii) support networks. Figure 7 presents the ITT impacts on five variables that explore these potential mechanisms. For self-efficacy, we used the Brief Parental Self-Efficacy Scale (BPSES) at the first follow-up but adapted to the Tool to Measure Parenting Self-Efficacy (TOPSE) at the second follow-up for more detailed questions relating to discipline and self-acceptance. For networks, we study parenting support and borrowing support networks.

Figure 7 shows positive impacts on caregivers' self-efficacy relating to self-acceptance at the second follow-up (0.21 SD, $p < 0.01$). We do not find statistically significant impacts on caregivers' parenting support or borrowing networks. The lack of significant impacts on support networks highlights that peer interactions during the weekly virtual sessions were likely limited.

The positive treatment effects and analysis of mechanisms suggest that the intervention provided caregivers with the necessary parenting tools and boosted their confidence in their own parenting skills. Taken together, our results suggest that the harsh behaviors of the caregivers could be explained, to a large extent, by a lack of information and preparedness. As Baker-Henningham and Francis (2018) have suggested, integrating new interventions that aim to train parents in alternative discipline strategies could help to improve the quality of parenting and, therefore, to reduce violent behaviors against the children.

[27] Responsiveness behaviors are based on sensitive, acceptance and warmth responses towards the child's actions (Landry et al., 2012).

8 Cost Effectiveness

Table 6 breaks down the fixed and variable costs (in 2021 USD) associated with each intervention component. The costs associated with sending SMS messages to caregivers in the treatment group over the duration of the intervention totaled USD 794.28. Dividing this across the 557 caregivers in the treatment group yields a per-caregiver cost of USD 1.43 (regardless of take-up). This highlights that the main component of our intervention was extremely cost-effective and scalable.

Fixed costs associated with the development of the KnowHub App were USD 30,104.51, with the majority of the costs associated with the filming of the videos for the App (USD 18,903.76). Fixed costs associated with the virtual sessions totaled USD 14,730.44, in large part owing to staff costs. The staff costs comprise consultant salaries but do not include salaries of the ECC officers since they were not provided with additional compensation for delivering the intervention. Variable costs associated with the virtual sessions were for data plans for caregivers to participate in the sessions. These costs were USD 19,235.25 in total.

9 Conclusion

The spectre of violence against young children is challenging. Worldwide, 300 million children aged 2-to-4 years are regularly subjected to violent discipline by their caregivers at home. In Jamaica, 85% of children 2-to-4 years experience violence at home and they are more likely to be subjected to physical punishment compared to older (5-to-14 years old) children. Despite violence undermining children's sense of self-worth and hindering their development, violence against children is often rationalized as necessary or inevitable and can be accepted when it is perpetrated by a "familiar face" (UNICEF, 2017).

The incidence of child maltreatment is more relevant during emergencies, such as pandemics. We evaluate the impact of the virtual Irie Homes Toolbox, a free online parent-training, violence-prevention program for Jamaican parents of children aged two-to-six years. We conduct an individual-level experiment during the lockdown measures due to COVID-19 in Jamaica. Our results show that the intervention improved caregivers' attitudes and violent behaviors against children in the short term. The program also reduced children's emotional problems after the program was completed. Notably, our results indicate that these effects of the intervention on caregivers' attitudes and violent behaviors against children persist after nine months of the first follow up. We document that the potential mechanisms driving these effects are improvements in parental self-efficacy and caregivers' mental well-being. Overall, the positive impacts of the virtual adaptation of the Irie Homes Toolbox support the call for broader implementation of interventions aiming to improve parenting literacy and self-efficacy.

From a policy perspective, the positive effects of the intervention on attitudes and behaviors related

to violent discipline against children provide evidence of the importance of developing and implementing parenting training for caregivers. Digital information campaigns that aim to build positive relationships between parent and child, prevent misbehavior, manage misbehavior, and emotional self-regulation help to increase the quality of parenting. Moreover, the combination of these learnings can help to reduce harmful parenting practices during the early ages of the children, diminishing exposition to violence at home (Baker-Henningham and Francis, 2018). Altogether, a more positive and healthy rearing will counteract the negative impacts of violent attitudes and environments such as risky behaviors and performance in school (Hamby et al., 2011; Carrell and Hoekstra, 2010; Hentges and Wang, 2018), increasing the long-term well-being of the children and the families.

References

Aguero, J. and Frisancho, V. (2021). Measuring violence against women with experimental methods. *Economic Development and Cultural Change*.

Amaral, S., Dinarte Diaz, L., Dominguez, P., and Perez-Vincent, S. M. (2021). Helping families help themselves? heterogeneous effects of a digital parenting program.

Amaral, S., Lelys, D.-D., Patricio, D., Steffanny, R., and Santiago, P.-V. (2022). Talk or text? evaluating response rates by remote survey method during covid-19. *Policy Research Working Paper Series*, 9999.

Anderson, M. L. (2008). Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training projects. *Journal of the American Statistical Association*, 103(484):1481–1495.

Asadullah, M. N., De Cao, E., Khatoon, F. Z., and Siddique, Z. (2021). Measuring gender attitudes using list experiments. *Journal of Population Economics*, 34(2):367–400.

Attanasio, O., Cattan, S., Fitzsimons, E., Meghir, C., and Rubio-Codina, M. (2020a). Estimating the production function for human capital: Results from a randomized controlled trial in Colombia. *American Economic Review*, 110(1):48–85.

Attanasio, O., Meghir, C., and Nix, E. (2020b). Human capital development and parental investment in India. *The Review of Economic Studies*, 87(6):2511–2541.

Attanasio, O. P., Fernández, C., Fitzsimons, E. O., Grantham-McGregor, S. M., Meghir, C., and Rubio-Codina, M. (2014). Using the infrastructure of a conditional cash transfer program to deliver a scalable integrated early child development program in colombia: cluster randomized controlled trial. *Bmj*, 349.

Baker-Henningham, H. and Francis, T. (2018). Parents' use of harsh punishment and young children's behaviour and achievement: a longitudinal study of jamaican children with conduct problems. *Global Mental Health*, 5.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American psychologist*, 37(2):122.

Bandura, A. and Wessels, S. (1994). Self-efficacy.

Baranov, V., Bhalotra, S., Biroli, P., and Maselko, J. (2020). Maternal depression, women's empowerment, and parental investment: Evidence from a randomized controlled trial. *American Economic Review*, 110(3):824–59.

Barrera, O., Macours, K., Premand, P., and Vakis, R. (2020). Texting parents about early child development. *World Bank Policy Research Working Paper*, (9492).

Belsky, J. (1984). The determinants of parenting: A process model. *Child development*, pages 83–96.

Belsky, J. and Jaffee, S. R. (2006). The multiple determinants of parenting.

Bendini, M. and Dinarte, L. (2020). Does maternal depression undermine childhood cognitive development? Evidence from the Young Lives Survey in Peru. *International Journal of Environmental Research and Public Health*, 17(19):7248.

Berry, J. O. and Jones, W. H. (1995). The parental stress scale: Initial psychometric evidence. *Journal of social and personal relationships*, 12(3):463–472.

Bloomfield, L. and Kendall, S. (2012). Parenting self-efficacy, parenting stress and child behaviour before and after a parenting programme. *Primary health care research & development*, 13(4):364–372.

Boden, J. M., Horwood, L. J., and Fergusson, D. M. (2007). Exposure to childhood sexual and physical abuse and subsequent educational achievement outcomes. *Child abuse & neglect*, 31(10):1101–1114.

Bruhn, M. and McKenzie, D. (2009). In pursuit of balance: Randomization in practice in development field experiments. *American Economic Journal: Applied Economics*, 1(4):200–232.

Carneiro, P. M., Galasso, E., Lopez Garcia, I. X., Bedregal, P., and Cordero, M. (2019). Parental beliefs, investments, and child development: Evidence from a large-scale experiment. *World Bank Policy Research Working Paper*, (8743).

Carrell, S. E. and Hoekstra, M. L. (2010). Externalities in the classroom: How children exposed to domestic violence affect everyone's kids. *American Economic Journal: Applied Economics*, 2(1):211–28.

Chorpita, B. F. and Daleiden, E. L. (2009). Mapping evidence-based treatments for children and adolescents: application of the distillation and matching model to 615 treatments from 322 randomized trials. *Journal of consulting and clinical psychology*, 77(3):566.

Cortes, K. E., Fricke, H. D., Loeb, S., Song, D. S., and York, B. N. (2019). When behavioral barriers are too high or low–how timing matters for parenting interventions. Technical report, National Bureau of Economic Research.

Crowne, D. P. and Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology*, 24(4):349.

Currie, J. and Spatz Widom, C. (2010). Long-term consequences of child abuse and neglect on adult economic well-being. *Child maltreatment*, 15(2):111–120.

Currie, J. and Tekin, E. (2012). Understanding the cycle childhood maltreatment and future crime. *Journal of Human Resources*, 47(2):509–549.

Dhar, D., Jain, T., and Jayachandran, S. (2022). Reshaping adolescents' gender attitudes: Evidence from a school-based experiment in india. *American Economic Review*, 112(3):899–927.

Doepke, M., Sorrenti, G., and Zilibotti, F. (2019). The economics of parenting. *Annual Review of Economics*, 11:55–84.

Donker, T., van Straten, A., Marks, I., and Cuijpers, P. (2011). Quick and easy self-rating of generalized anxiety disorder: validity of the dutch web-based gad-7, gad-2 and gad-si. *Psychiatry research*, 188(1):58–64.

Doss, C., Fahle, E. M., Loeb, S., and York, B. N. (2019). More than just a nudge supporting kindergarten parents with differentiated and personalized text messages. *Journal of Human Resources*, 54(3):567–603.

Doyle Jr, J. J. and Aizer, A. (2018). Economics of child protection: Maltreatment, foster care, and intimate partner violence. *Annual review of economics*, 10:87–108.

Francis, T. and Baker-Henningham, H. (2020). Design and implementation of the irie homes toolbox: a violence prevention, early childhood, parenting program. *Frontiers in public health*, 8:702.

Francis, T., Dinarte-Diaz, L., Powers, S., and Baker-Henningham, H. (2022). The virtual irie homes toolbox: Development and implementation.

Fry, D., Fang, X., Elliott, S., Casey, T., Zheng, X., Li, J., Florian, L., and McCluskey, G. (2018). The relationships between violence in childhood and educational outcomes: A global systematic review and meta-analysis. *Child abuse & neglect*, 75:6–28.

Garland, A. F., Hawley, K. M., Brookman-Frazee, L., and Hurlburt, M. S. (2008). Identifying common elements of evidence-based psychosocial treatments for children's disruptive behavior problems. *Journal of the American Academy of Child & Adolescent Psychiatry*, 47(5):505–514.

Hamadani, J. D., Mehrin, S. F., Tofail, F., Hasan, M. I., Huda, S. N., Baker-Henningham, H., Ridout, D., and Grantham-McGregor, S. (2019). Integrating an early childhood development programme into bangladeshi primary health-care services: an open-label, cluster-randomised controlled trial. *The Lancet Global Health*, 7(3):e366–e375.

Hamby, S., Finkelhor, D., Turner, H., and Ormrod, R. (2011). *Children's Exposure to Intimate Partner Violence and Other Family Violence*. Washington, DC: U.S. Department of Justice.

Heckman, J. J., Liu, B., Lu, M., and Zhou, J. (2020). The impacts of a prototypical home visiting program on child skills. *NBER Working Paper*, (w27356).

Hentges, R. F. and Wang, M.-T. (2018). Gender differences in the developmental cascade from harsh parenting to educational attainment: An evolutionary perspective. *Child development*, 89(2):397–413.

Hoover-Dempsey, K. V., Walker, J. M., Sandler, H. M., Whetsel, D., Green, C. L., Wilkins, A. S., and Closson, K. (2005). Why do parents become involved? research findings and implications. *The elementary school journal*, 106(2):105–130.

Hurwitz, L. B., Lauricella, A. R., Hanson, A., Raden, A., and Wartella, E. (2015). Supporting head start parents: impact of a text message intervention on parent–child activity engagement. *Early Child Development and Care*, 185(9):1373–1389.

Izzo, C., Weiss, L., Shanahan, T., and Rodriguez-Brown, F. (2000). Parental self-efficacy and social support as predictors of parenting practices and children's socioemotional adjustment in mexican immigrant families. *Journal of Prevention & Intervention in the Community*, 20(1-2):197–213.

Jeong, J., Obradović, J., Rasheed, M., McCoy, D. C., Fink, G., and Yousafzai, A. K. (2019). Maternal and paternal stimulation: Mediators of parenting intervention effects on preschoolers' development. *Journal of Applied Developmental Psychology*, 60:105–118.

Jeong, J., Pitchik, H. O., and Yousafzai, A. K. (2018). Stimulation interventions and parenting in low-and middle-income countries: a meta-analysis. *Pediatrics*, 141(4).

Jones, T. L. and Prinz, R. J. (2005). Potential roles of parental self-efficacy in parent and child adjustment: A review. *Clinical psychology review*, 25(3):341–363.

JPAL (2020). Encouraging early childhood stimulation from parents and caregivers to improve child development. Technical report.

Kroenke, K., Spitzer, R. L., and Williams, J. B. (2003). The patient health questionnaire-2: validity of a two-item depression screener. *Medical care*, pages 1284–1292.

Landry, S. H., Smith, K. E., Swank, P. R., Assel, M. A., and Vellet, S. (2001). Does early responsive parenting have a special importance for children's development or is consistency across early childhood necessary? *Developmental psychology*, 37(3):387.

Landry, S. H., Smith, K. E., Swank, P. R., Zucker, T., Crawford, A. D., and Solari, E. F. (2012). The effects of a responsive parenting intervention on parent–child interactions during shared book reading. *Developmental psychology*, 48(4):969.

Lansford, J. E., Dodge, K. A., Pettit, G. S., Bates, J. E., Crozier, J., and Kaplow, J. (2002). A 12-year prospective study of the long-term effects of early child physical maltreatment on psychological, behavioral, and academic problems in adolescence. *Archives of pediatrics & adolescent medicine*, 156(8):824–830.

Lee, D. S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *The Review of Economic Studies*, 76(3):1071–1102.

Levere, M., Acharya, G., and Bharadwaj, P. (2016). The role of information and cash transfers on early childhood development: evidence from nepal. Technical report, National Bureau of Economic Research.

Linhares, M. B. M., Altafim, E. R., de Oliveira, R. C., and Gaspardo, C. M. (2022). A personalized remote video-feedback universal parenting program: a randomized controlled trial. *Psychosocial Intervention*, 31(1):21–32.

Macours, K., Schady, N., and Vakis, R. (2012). Cash transfers, behavioral changes, and cognitive development in early childhood: evidence from a randomized experiment. *American Economic Journal: Applied Economics*, 4(2):247–73.

Mehrin, S. F., Hasan, M. I., Tofail, F., Shiraji, S., Ridout, D., Grantham-McGregor, S., Hamadani, J. D., and Baker-Henningham, H. (2022). Integrating a group-based, early childhood parenting intervention into primary health care services in rural bangladesh: A cluster-randomised controlled trial. *Frontiers in Pediatrics*.

Mersky, J. P. and Topitzes, J. (2010). Comparing early adult outcomes of maltreated and non-maltreated children: A prospective longitudinal investigation. *Children and Youth Services Review*, 32(8):1086–1096.

Michl-Petzing, L. C., Handley, E. D., Sturge-Apple, M., Cicchetti, D., and Toth, S. L. (2019). Re-examining the "cycle of abuse": Parenting determinants among previously maltreated, low-income mothers. *Journal of family psychology*, 33(6):742.

Olivetti, C. and Petrongolo, B. (2017). The economic consequences of family policies: lessons from a century of legislation in high-income countries. *Journal of Economic Perspectives*, 31(1):205–30.

Paxson, C. and Schady, N. (2010). Does money matter? the effects of cash transfers on child development in rural ecuador. *Economic development and cultural change*, 59(1):187–229.

Persson, P. and Rossin-Slater, M. (2018). Family ruptures, stress, and the mental health of the next generation. *American Economic Review*, 108(4-5):1214–52.

Sviatschi, M. M. (2022). Making a narco: Childhood exposure to illegal labor markets and criminal life paths. *Econometrica*, 90(4):1835–1878.

Taraban, L. and Shaw, D. S. (2018). Parenting in context: Revisiting belsky's classic process of parenting model in early childhood. *Developmental Review*, 48:55–81.

UNICEF (2017). A familiar face: Violence in the lives of children and adolescents. *New York, NY: UNICEF*.

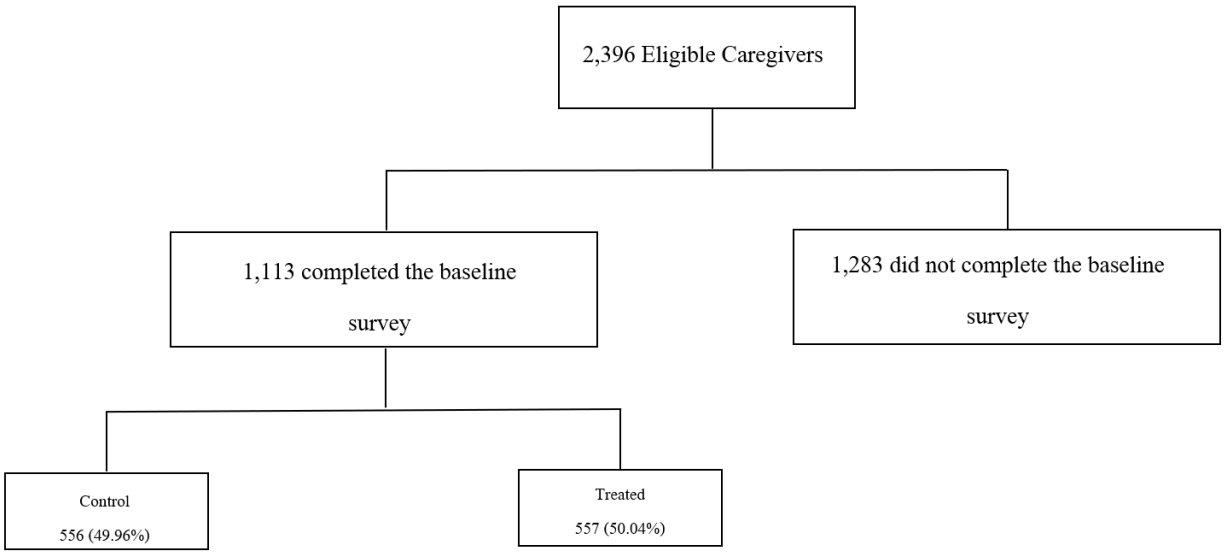
Widen, S. C., Orozco, M., Horng, E. L., and Loeb, S. (2020). Reaching unconnected caregivers: Using a text-message education program to better understand how to support informal caregivers roles in child development. *Journal of Early Childhood Research*, 18(1):29–43.

York, B. N., Loeb, S., and Doss, C. (2019). One step at a time the effects of an early literacy text-messaging program for parents of preschoolers. *Journal of Human Resources*, 54(3):537–566.

Tables and Figures

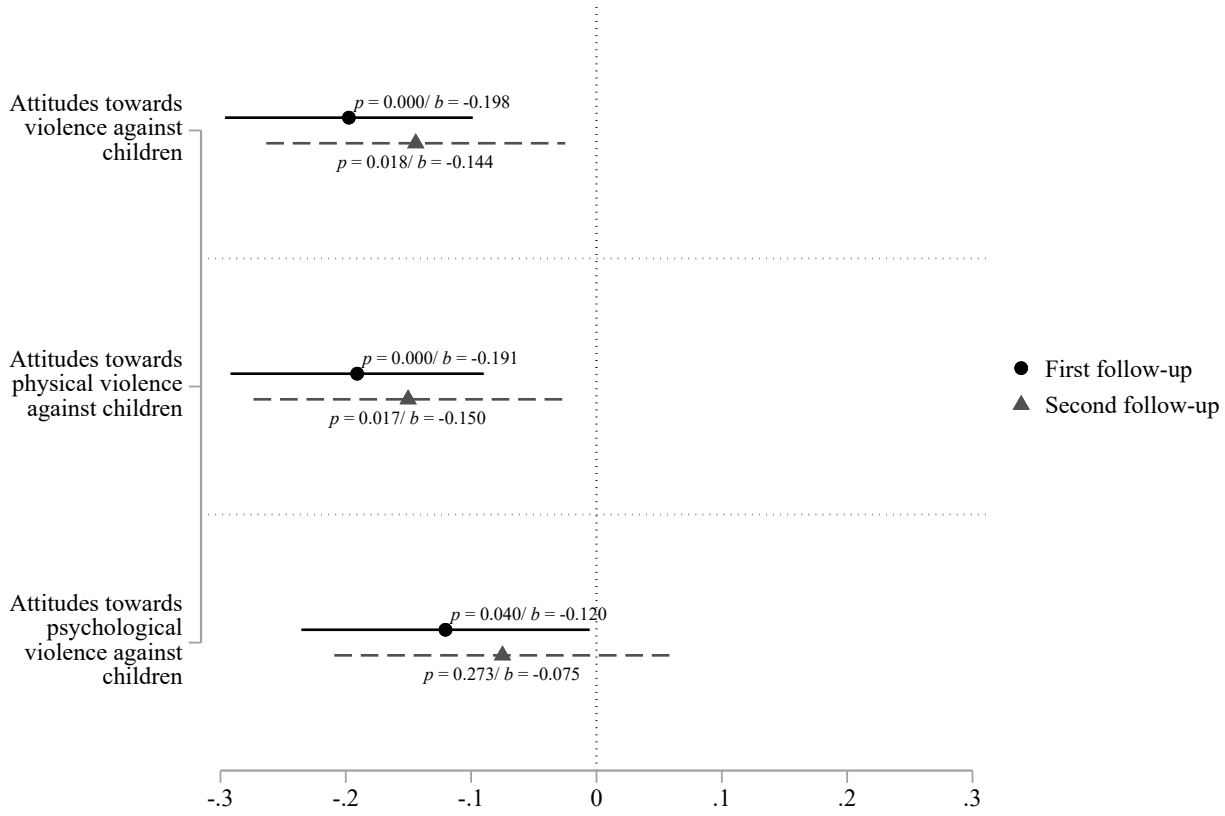
Figures

Figure 1: Experimental Design



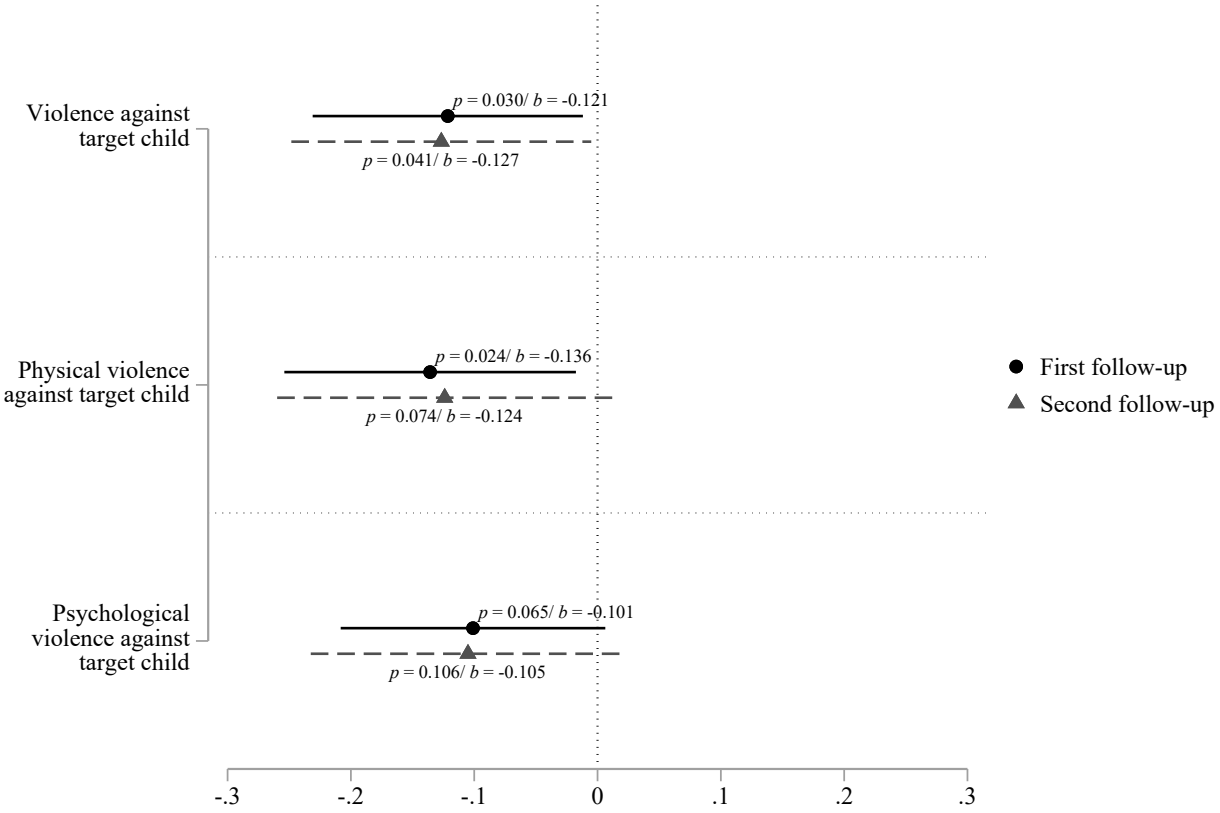
Notes: This figure summarizes the experimental design of the study. 1,283 enrolled individuals did not complete the baseline survey for several reasons, including they did not provide a correct phone number; we were unable to reach them after the maximum number of attempts determined in the ethics protocol; they changed their mind and decided not to participate in the study, among others.

Figure 2: ITT Impacts on Caregiver Attitudes Related to Violence Towards Their Child (Indices)



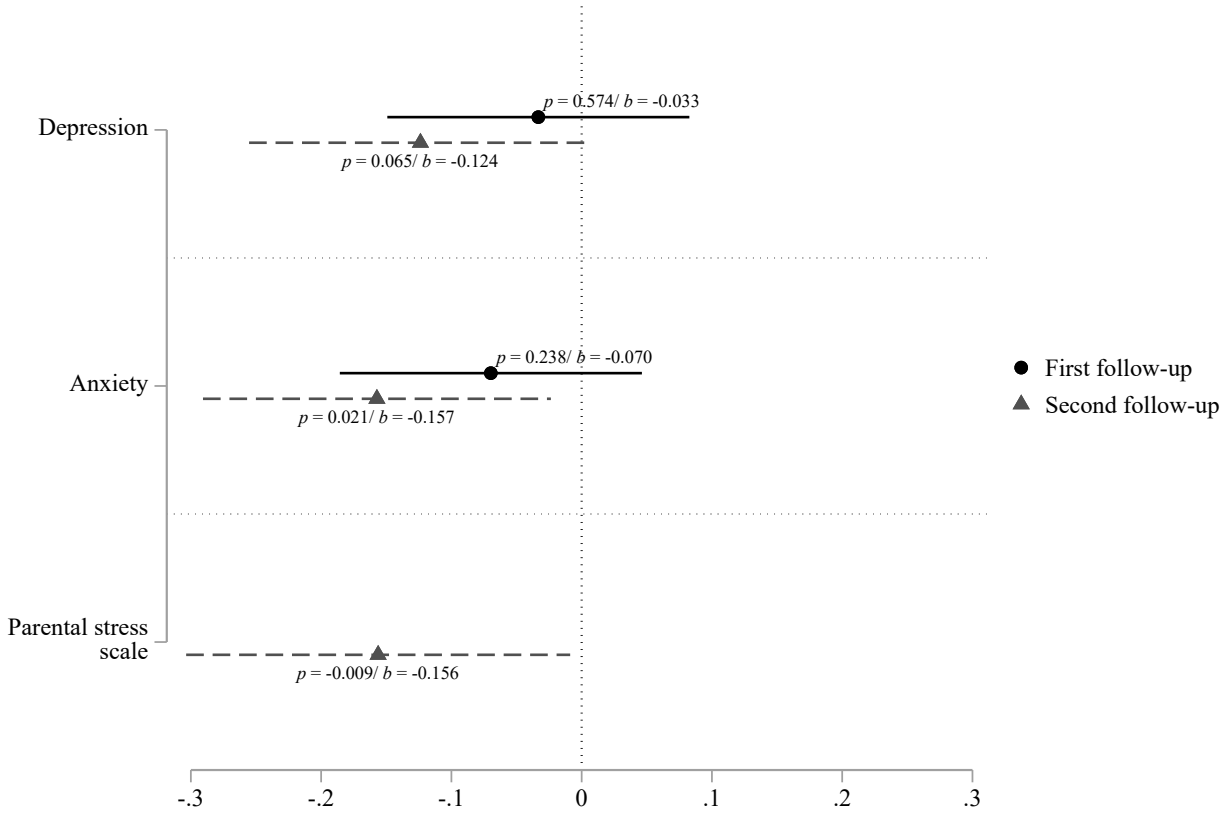
Notes: This figure presents OLS estimates of the short and medium-term intent-to-treat impacts on caregivers' attitudes relating to violence against children (indices). Each outcome consists of a standardized index estimated following Anderson (2008) and standardized relative to the control group. For a detailed description of the indices, see Section 3.3.2. The black circles and corresponding solid lines represent the point estimates and 95% confidence intervals from the first follow-up survey. The gray triangles and corresponding dashed lines represent the point estimates and 95% confidence intervals from the second follow-up survey. Standard deviation units are used for the x-axis.

Figure 3: ITT Impacts on Caregiver Behaviors Related to Violence Toward the Target Child (Indices)



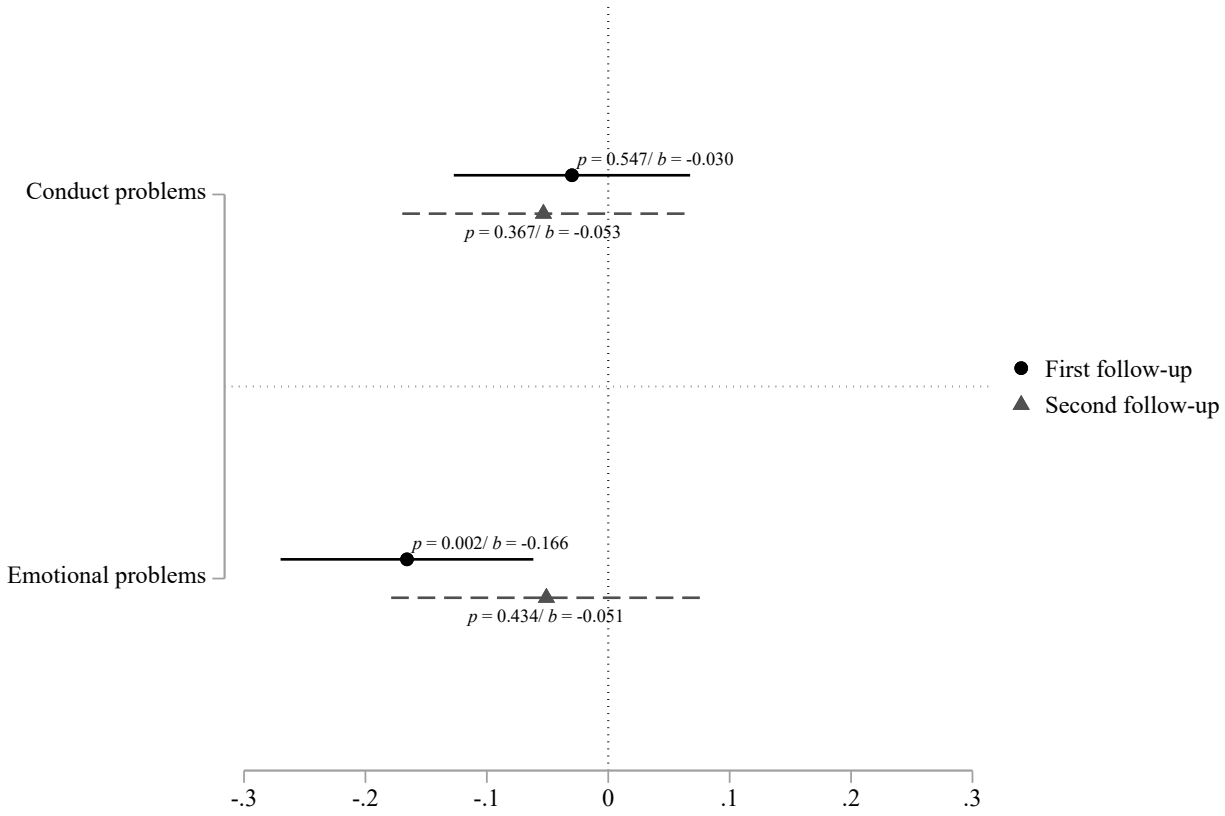
Notes: This figure presents OLS estimates of the short and medium-term intent-to-treat impacts on caregivers' behaviors (indices) relating to violence against the target child. Each outcome consists of a standardized index estimated following Anderson (2008) and standardized relative to the control group. For a detailed description of the indices, see Section 3.3.2. The black circles and corresponding solid lines represent the point estimates and 95% confidence intervals from the first follow-up survey. The gray triangles and corresponding dashed lines represent the point estimates and 95% confidence intervals from the second follow-up survey. Standard deviation units are used for the x-axis.

Figure 4: ITT Impacts on Caregivers' Depression, Anxiety, & Stress (Indices)



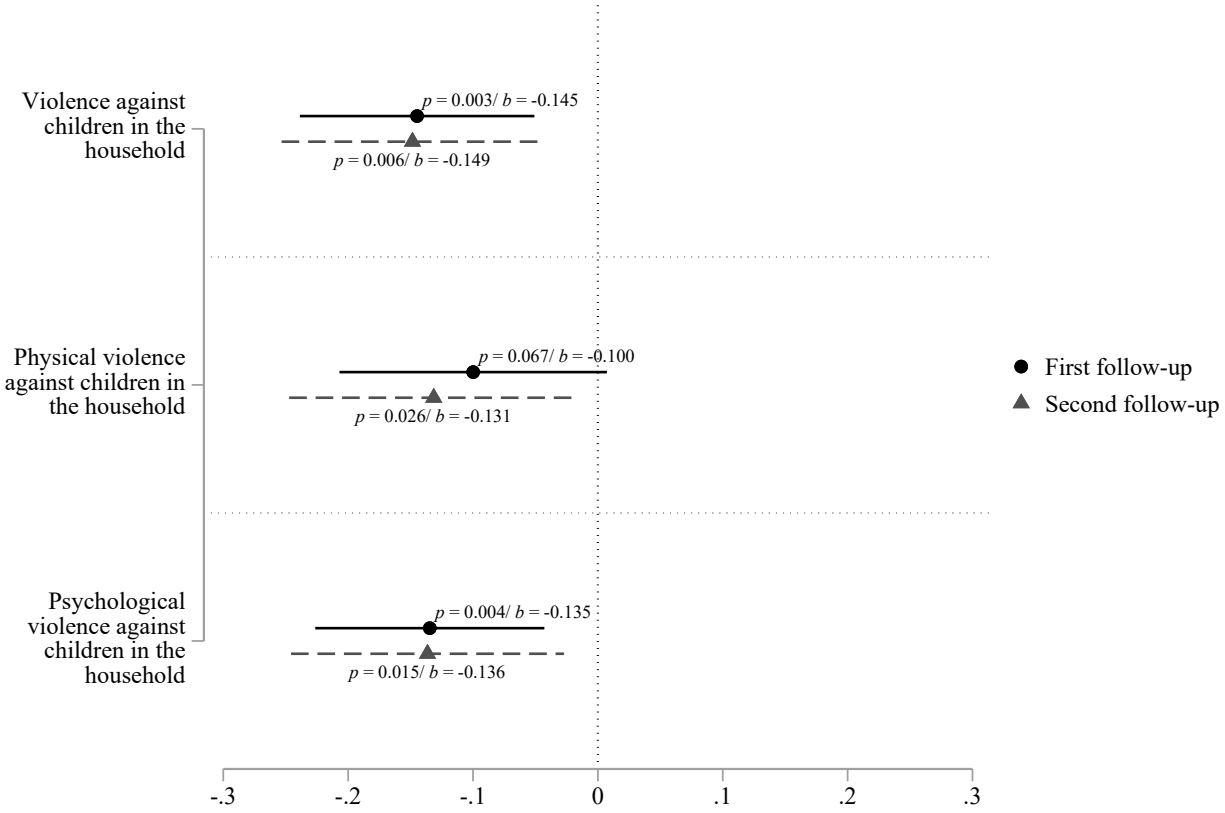
Notes: This figure presents OLS estimates of the short and medium-term intent-to-treat impacts of the intervention on caregivers' depression (measured using PHQ-2 at both follow-ups), anxiety (measured using GAD-2 at both follow-ups), and parental stress (measured using PSS-18 only at second follow-up). Each outcome consists of a standardized index estimated following Anderson (2008) and standardized relative to the control group. For a detailed description, see Section 3.3.2. The black circles and corresponding solid lines represent the point estimates and 95% confidence intervals from the first follow-up survey. The gray triangles and corresponding dashed lines represent the point estimates and 95% confidence intervals from the second follow-up survey. Standard deviation units are used for the x-axis.

Figure 5: ITT Impacts on Child Conduct and Emotional Problems (Indices)



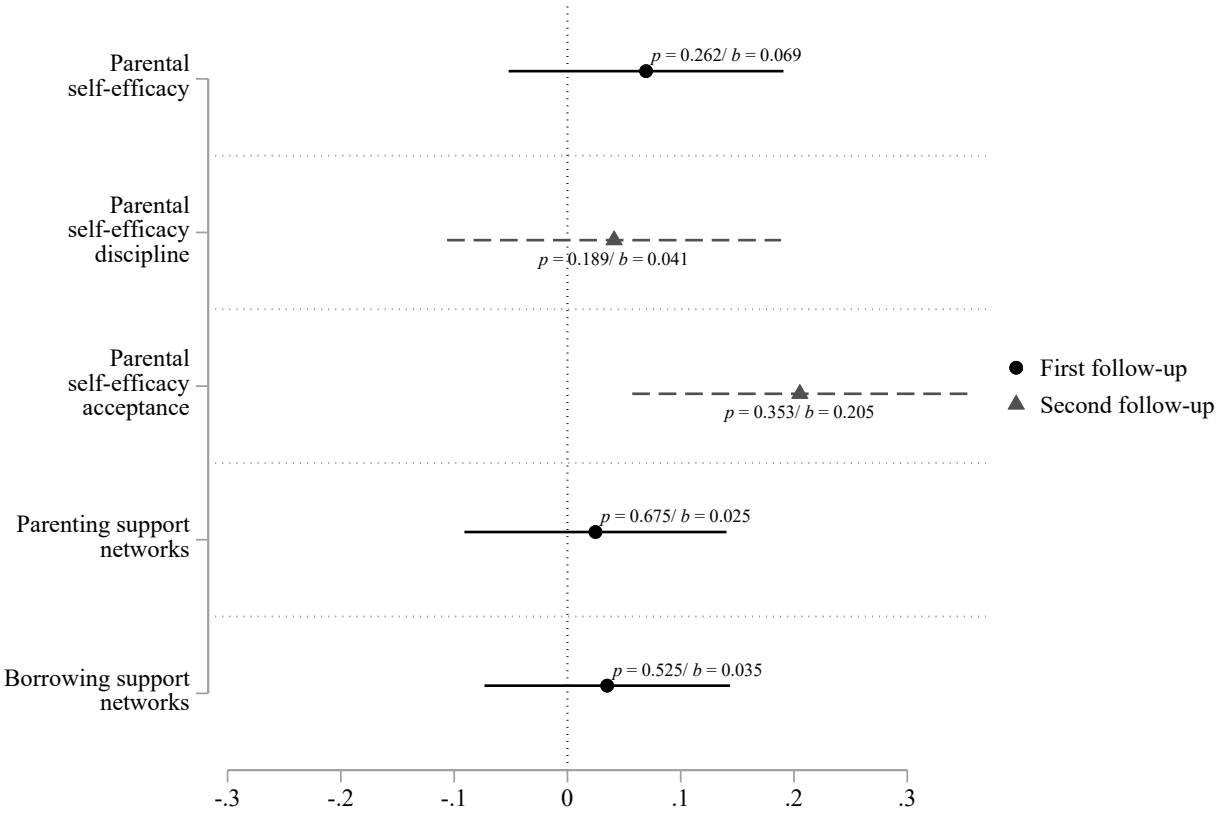
Notes: This figure presents OLS estimates of the short and medium-term intent-to-treat impacts of the intervention on the target child's conduct and emotional problems. Conduct and emotional problems were measured using caregivers' responses to the Strengths and Difficulties Questionnaire (SDQ) instrument. Each outcome consists of a standardized index estimated following Anderson (2008) and standardized relative to the control group. For a detailed description, see Section 3.3.2. The black circles and corresponding solid lines represent the point estimates and 95% confidence intervals from the first follow-up survey. The gray triangles and corresponding dashed lines represent the point estimates and 95% confidence intervals from the second follow-up survey. Standard deviation units are used for the x-axis.

Figure 6: ITT Impacts on Caregiver Behaviors Related to Violence Toward Children (Indices)



Notes: This figure presents OLS estimates of the short and medium-term intent-to-treat impacts on caregivers' behaviors (indices) relating to violence against the target child and the eldest child (if any, aged 7-12) in the household. Each outcome consists of a standardized index estimated following Anderson (2008) and standardized relative to the control group. For a detailed description of the indices, see Section 3.3.2. The black circles and corresponding solid lines represent the point estimates and 95% confidence intervals from the first follow-up survey. The gray triangles and corresponding dashed lines represent the point estimates and 95% confidence intervals from the second follow-up survey. Standard deviation units are used for the x-axis.

Figure 7: Mechanisms (Indices)



Notes: This figure presents OLS estimates of the short and medium-term intent-to-treat impacts on five variables that explore potential mechanisms through which the intervention improved the caregivers' self-efficacy and support networks. For self-efficacy, we used the Brief Parental Self-Efficacy Scale (BPSES) at the first follow-up but adapted to the Tool to Measure Parenting Self-Efficacy (TOPSE) at the second follow-up for more detailed questions relating to discipline and self-acceptance. For networks, we examined parenting support networks and borrowing support networks. Each outcome consists of a standardized index estimated following Anderson (2008) and standardized relative to the control group. For a detailed description, see Section 3.3.2. The black circles and corresponding solid lines represent the point estimates and 95% confidence intervals from the first follow-up survey. The gray triangles and corresponding dashed lines represent the point estimates and 95% confidence intervals from the second follow-up survey. Standard deviation units are used for the x-axis.

Tables

Table 1: Summary Statistics by Group and Balance Tests

	Control			Treatment			P-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	N	Mean	SE	N	Mean	SE	(2) - (5)
Panel A. Caregiver's characteristics							
Age	556	32.894	0.333	557	32.553	0.324	0.443
Gender - female (%)	556	0.995	0.004	557	0.991	0.006	0.634
Education level completed	556	14.243	0.124	557	14.471	0.124	0.166
Marital status - married (%)	380	0.362	0.027	402	0.336	0.025	0.450
Employed (%)	476	0.778	0.020	499	0.788	0.020	0.700
Income in the past month (USD)	429	881.106	61.796	448	829.315	51.531	0.489
Household total size	556	4.585	0.086	557	4.624	0.090	0.734
Total number of children 17 or below	556	1.973	0.046	557	1.896	0.048	0.213
Food insecurity - Eating less (%)	556	0.225	0.018	557	0.252	0.019	0.266
Panel B. Child's characteristics'							
Target child 2-6 years old - Age	556	4.164	0.065	557	4.063	0.064	0.236
Target child 2-6 years old - Female (%)	556	0.486	0.023	557	0.501	0.023	0.623
Panel C. Primary Outcomes							
Violence against target child (index)	556	0.065	0.045	557	0.046	0.041	0.736
Avg. # of days with discipline attitudes - target child	556	1.425	0.055	557	1.437	0.053	0.867
Violence against other child in the hh (index)	556	0.021	0.030	557	0.014	0.029	0.853
Avg. # of days with discipline attitudes - oldest child	556	1.092	0.037	557	1.087	0.037	0.916
Attit. to viol against children (index)	556	-0.045	0.043	557	0.004	0.045	0.417
(%) agreement with violent attitudes	556	16.863	0.928	557	17.753	0.956	0.492
Panel D. Secondary Outcomes							
Conduct problems (index)	556	0.039	0.044	557	0.028	0.046	0.854
(%) Conduct problems	556	42.566	1.201	557	42.874	1.230	0.849
Emotional problems (index)	556	0.011	0.045	557	-0.004	0.042	0.801
(%) Emotional problems	556	25.985	1.149	557	26.312	1.095	0.824
Depression (index)	556	0.038	0.045	557	0.032	0.044	0.912
(%) High depression	556	20.568	1.804	557	18.748	1.773	0.440
Anxiety (index)	556	0.018	0.045	557	0.088	0.044	0.242
(%) High anxiety	556	12.883	1.544	557	16.598	1.676	0.084
Caregiver involvement in play and learning activ.(index)	556	0.047	0.045	557	0.087	0.043	0.492
Average number of good parenting practices	556	4.804	0.073	557	4.851	0.069	0.623
Panel E. Mechanisms							
Borrowing money support (index)	556	2.166	0.112	557	2.200	0.106	0.825
Parenting issues support (index)	556	2.519	0.122	557	2.705	0.121	0.247
Belong to parent support group (%)	556	0.145	0.016	557	0.152	0.016	0.761
F-test of joint significance (p-value)							0.947

Notes: This table shows average characteristics at baseline for the study participants assigned to the treatment and control groups. Columns (1) and (4) present the number of caregivers in the control and treatment groups; (2) and (5) the mean; (3) and (6) standard errors, and (7) the p-value associated with the hypothesis of the mean values across both groups being the same. We imputed the mean to have consistent sample sizes in the following variables. **Panel A:** Food insecurity. **Panel C:** Violence against target child, Avg. # of days with discipline attitudes - target child, Violence against other children in the household, Avg. # of days with discipline attitudes - oldest child, Attit. to violence against children, (%) agreement with violent attitudes. **Panel E:** Borrowing money support, Parenting issues support, Belonging to a parent support group, Caregiver involvement in play and learning activities, Average number of good parenting practices.

Table 2: First Stage: Take-up of Intervention

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A. SMS					
Received SMS (%)	91.38	28.09	0	1	499
Read SMS if received (%)	96.94	17.23	0	1	458
Found the SMS useful if read (%)	98.20	13.32	0	1	444
Panel B. App usage					
Number of sessions accessed	1.04	1.85	0	10	557
Total time in sessions (mins)	6.94	15.58	0	75	557
Panel C. Virtual sessions					
Number of sessions attended	4.55	3.44	0	10	557

Notes: This table shows descriptive statistics for the take-up of relevant outcomes for each of the three components of the intervention. The table uses survey data on reception of the SMS messages, App usage data from the phone company, and ECC officer reports of attendance at the virtual sessions.

Table 3: First Stage: ITT Impacts on Learning

	Information module								
	(1) Praising children helps	(2) Imp. for parents to play w/ child	(3) Clear instructions help	(4) Understand why child misbehaves	(5) Calm down before disciplining	(6) Withdraw attention from childś whining	(7) Redirect rather than reprimand	(8) Consequences and timeout appropriate	(9) Information module (index)
Panel A: First follow-up									
Treatment	0.189***	0.059	0.224***	0.063*	-0.011	0.624***	0.278***	0.080**	0.525***
	(0.05)	(0.04)	(0.05)	(0.04)	(0.04)	(0.07)	(0.05)	(0.04)	(0.07)
Observations	978	979	978	974	974	971	973	971	979
R^2	0.02	0.00	0.03	0.00	0.00	0.08	0.03	0.01	0.05
Control mean	4.13	4.27	3.98	4.29	4.38	2.73	3.74	4.16	-0.00
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Panel B: Second follow-up									
Treatment	0.104**	0.039	0.248***	-0.058	-0.026	0.480***	0.165***	0.071	0.394***
	(0.05)	(0.04)	(0.06)	(0.04)	(0.04)	(0.08)	(0.06)	(0.04)	(0.08)
Observations	698	699	699	699	698	691	691	697	699
R^2	0.01	0.00	0.03	0.00	0.00	0.05	0.01	0.01	0.04
Control mean	4.196	4.290	3.989	4.408	4.450	2.743	3.810	4.164	0.000

Notes: This table shows the intention-to-treat estimates of the treatment on caregivers' learning outcomes. Panel A presents the effects corresponding to the first follow-up. Panel B shows the medium treatment effect corresponding to the second follow-up, nine months after the original intervention. Treatment is a dummy variable equal to 1 if the caregiver is assigned to the treatment group and zero otherwise. Columns 1 to 8 show the impact of treatment over eight statements relating to parenting practices. Each statement is designed to evaluate the understanding of the four key concepts of the intervention. All outcome variables are Likert-scale variables ranging from 1 (Strongly disagree) to 5 (Strongly agree). Respondents were asked about the degree to which they agreed with each statement. Column 9 presents the ITT impacts of the treatment on the information index, which aggregates the eight outcome statements as described in Section 4. Heteroskedasticity-robust standard errors are reported in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4: Treatment Impacts Controlling for Social Desirability Bias (Indices)

	Primary hypotheses				
	(1)	(2)	(3)	(4)	(5)
	SDB	Attitudes towards violence against children	Violence against target child	Violence against children in the household	Caregiver involvement in play and learning activ.
Panel A: First follow-up					
Treatment		-0.327***	-0.157**	-0.177***	-0.012
		(0.06)	(0.07)	(0.06)	(0.06)
SDB (index)		-0.023	-0.032	-0.061**	0.036
		(0.03)	(0.03)	(0.03)	(0.03)
Observations		674	657	889	661
R^2		0.26	0.27	0.26	0.23
Control Mean		0.075	0.053	0.048	0.080
Panel B: Second follow-up					
Treatment	-0.016	-0.145**	-0.132**	-0.151***	0.005
	(0.08)	(0.06)	(0.06)	(0.05)	(0.07)
SDB (index)		-0.047	-0.019	-0.020	0.065*
		(0.03)	(0.03)	(0.03)	(0.04)
Observations	700	696	677	909	672
R^2	0.01	0.26	0.18	0.18	0.24
Control Mean	0.000	-0.000	0.001	-0.010	-0.004

Notes: Column (1) reproduces the result from Table A13, Panel B, Column (7). Columns (2)-(5) of this table show our results on caregiver attitudes and behaviors controlling for Social Desirability Bias (index). The term "Control Mean" describes the mean of the outcome for the control group as measured at second follow-up. The sample size in each specification varies according to the number of observations available for each outcome. Heteroskedasticity-robust standard errors are reported in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5: Heterogeneity by Social Desirability Bias

	Primary hypotheses			
	(1) Attitudes to violence against children	(2) Violence against target child	(3) Violence against children in the household	(4) Caregiver involvement in play and learning activ.
Panel A: First follow-up				
Treatment	-0.307***	-0.115	-0.188**	0.004
	(0.08)	(0.11)	(0.09)	(0.09)
High SDB score	-0.030	-0.093	-0.155*	0.161*
	(0.10)	(0.10)	(0.09)	(0.09)
Treatment × High SDB score	-0.039	-0.090	0.019	-0.031
	(0.12)	(0.14)	(0.12)	(0.13)
Treat + Treat × High SDB score	-0.346***	-0.205**	-0.169**	-0.027
	(0.09)	(0.09)	(0.08)	(0.09)
Observations	674	657	889	661
Panel B: Second follow-up				
Treatment	-0.243***	-0.128	-0.165**	0.006
	(0.09)	(0.09)	(0.08)	(0.10)
High SDB score	-0.223**	-0.138	-0.138	0.114
	(0.09)	(0.10)	(0.09)	(0.10)
Treatment × High SDB score	0.198*	-0.009	0.029	-0.005
	(0.12)	(0.13)	(0.11)	(0.14)
	(1)	(2)	(3)	(4)
Treat + Treat × High SDB score	-0.045	-0.138*	-0.137*	0.001
	(0.08)	(0.08)	(0.07)	(0.09)
Observations	696	677	909	672

Notes: This table presents treatment heterogeneity of our results on caregiver attitudes and behaviors by Social Desirability Bias (index). "High SDB score" is a dummy variable that takes the value 1 if the SDB score was above the median SDB score for the sample, and 0 otherwise. The SDB score was only measured at second follow-up; this measure is used to examine heterogeneity in both the first and second follow-ups. The term "Treat + Treat x High SDB score" denotes the total effect of the treatment for those in the treatment group with above-median SDB scores. The sample size in each specification varies according to the number of observations available for each outcome. Heteroskedasticity-robust standard errors are reported in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Intervention Costs

Cost Category	Fixed Costs	Variable Costs	Total (USD)	Total per Caregiver Targeted (USD)
SMS	0.00	794.28	**794.28**	**1.43**
Costs to send SMSes	**0.00**	**794.28**		
App	**30,194.51**	**0.00**	**30,194.51**	**54.21**
Consultant and staff costs	3,562.75	0.00		
Costs associated with filming videos for the App	18,903.76	0.00		
KnowHub App development costs	7,728.00	0.00		
Virtual sessions	**14,730.44**	**19,235.25**	**33,965.69**	**60.98**
Consultant and staff costs	10,688.25	0.00		
Training materials	404.69	0.00		
Data plans to ECC officers to administer sessions	3,637.50	0.00		
Data plans to caregivers to participate in sessions	0.00	19,235.25		
Total (USD)	**44,924.95**	**20,029.53**	**64,954.48**	**116.61**
Total per Caregiver Targeted (USD)	**80.66**	**35.96**	**116.61**	

Notes: This table presents fixed and variable costs (in 2021 USD) associated with each intervention component. Costs were divided across 557 treatment group caregivers to compute the costs per caregiver targeted. For costs incurred in Jamaican Dollars, the following exchange rate was used: 1 USD = 154.94 JMD.

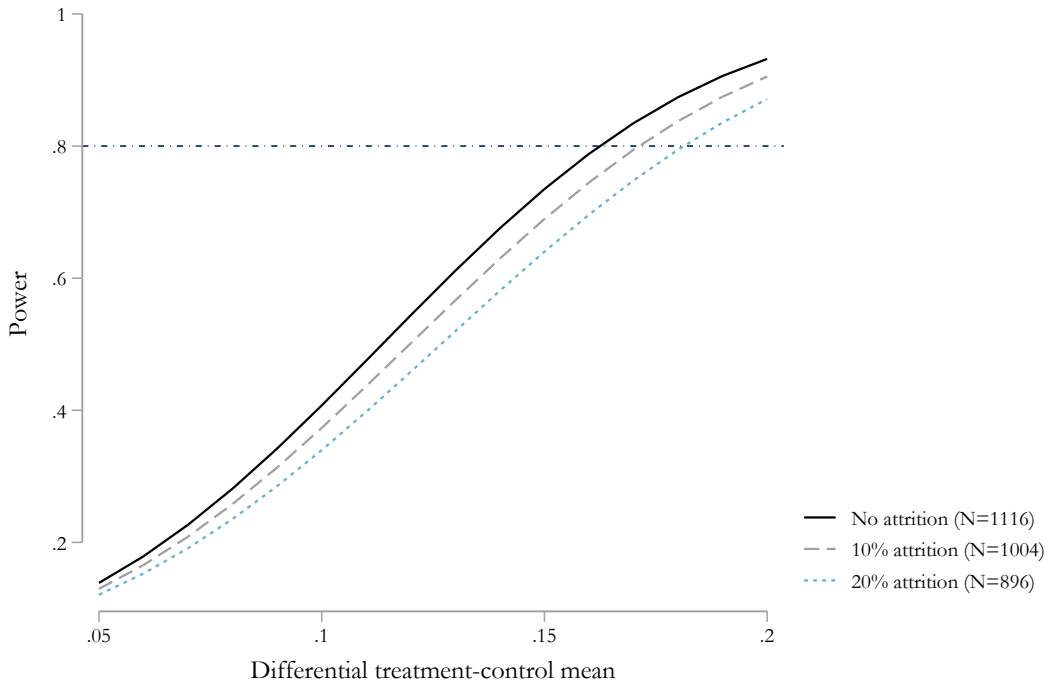
Appendix

For Online Publication Only

A1 Power Calculations

Figure A1 reports results of the power analysis. Assuming an α of 0.05, power of 80%, correlation between baseline and first follow-up (ANCOVA model) of 0.25, and 20% of attrition between baseline and first follow-up, we are powered to detect intention to treatment effects of approximately 0.18 standard deviations with respect to the control group. Besides, in order to illustrate the behaviour of our power simulations, Figure A1 also shows the relationship between power and minimum detectable effects for two additional scenarios: (i) no attrition (e.g. baseline full sample) and (ii) 10% attrition. In the first scenario (solid line-no attrition), with a power of 80% we are able to detect MDEs of 0.163. In a scenario of low attrition (10%), with 80% power we can detect MDEs of 0.173.

Figure A1: Power Analysis



Note: This figure reports results of the power analysis. We have assumed $\alpha = 0.05$ and correlation between baseline and first follow-up (ANCOVA model) $= 0.25$.

A2 Components of Indices

In this section we describe the components/questions of each index.

A2.1 Parental Support Activities

In this section, instructions were: *Please write down the number of days indicated by the respondent. If the respondent provides a range (for example, 2-3 days a week), please record the lowest number. Please ask for the number of days first, followed by minutes (do not ask both questions together)*. The reference period was **in the past 7 days**.

For how many days/minutes...

- Did you play with him/her?
- Did you read books or looked at picture books with him/her?
- Did you tell stories or sang songs to him/her?
- Did you named, counted, or drew things for or with him/her?
- Did you praise him/her when s/he did something good?

A2.2 Strengths and Difficulties

A2.2.1 Emotional Strengths and Difficulties

Questions drawn from the S&D questionnaire. Response options were: 0 - Not true, 1 - Somewhat true, 2 - Certainly true. The instruction was: *I have some more descriptions about your child's behaviour. After each of these statements I want you to think about your child's behaviour over the past month and tell me if the statement is certainly true about your child, somewhat true or not true. There are no wrong or right answers just try and answer as best as you can.*

- He/she often complains of headaches, stomach-aches, or sickness
- He/she has many worries, often seems worried
- He/she is often unhappy, downhearted, or tearful
- He/she is nervous or clingy in new situations, easily loses confidence
- He/she has many fears, is easily scared

A2.2.2 Conduct Strengths and Difficulties

Same structure as in previous section.

- He/she often has temper tantrums or hot tempers
- He/she is generally obedient, usually does what adults ask/request
- He/she often fights with other children or bullies them
- He/she is often argumentative with adults
- He/she is often spiteful to others

A2.3 Target Child and Eldest Child Discipline Indices

Same set of questions for both age groups. Questions asked for how many days in the past week did the parent do the action stated.

- Shouted, yelled, or screamed at him/her?
- Said you would send him/her away?
- Hit him/her on the bottom, hand, arm, or leg with your bare hand?
- Threatened to hit him/her but not actually done it?
- Hit him/her on the bottom, hand, arm, or leg with something like a belt, hairbrush, stick, or some other hard object?

A2.4 Attitudes Towards Violence Against the Child

Yes/no questions (agree or disagree).

- In order to bring up, raise, or educate a child properly, the child needs to be physically punished.
- A good parent slaps their child when they misbehave.
- When a child is beaten, he/she will stop doing the unwanted behavior.
- Shouting and yelling makes the child more obedient.
- Shouting, yelling, and threatening to slap will harm the child

A2.5 Depression Questions

Likert-scale variables. Instruction: *I will read you each statement and please indicate to me how often have you been bothered by any of the following problems over the last two weeks. Remember that there are no right or wrong answers. For each item, please indicate if you were bothered nearly every day, more than half the days, several days, or not at all*. Time reference period was **in the last two weeks**. Responses: 0 - Not at all, 1 - Several days, 2 - More than half the days, 3 - Nearly every day.

- Have you been feeling little interest or pleasure in doing things?
- Have you been feeling down, depressed, or hopeless?

A2.6 Anxiety Questions

Same as in previous section.

- Have you been feeling nervous, anxious or on edge?
- Have you not been able to stop or control worrying?

A2.7 Social Networks Indices

Social network variables were divided in the following groups:

A2.7.1 Borrowing Money Support

Instruction was: *How many people could you go to if you needed to borrow JMD 5,000? Please indicate separately for friends, family, and professionals. Professionals here could include bank officers and other moneylenders.*

- Friends: (type number)
- Family: (type number)
- Professionals: (type number)

A2.7.2 Parenting Issues Support

Instruction was: *How many people could you go to if you wanted to talk about issues relating to parenting and child rearing? Please indicate separately for friends, family, and professionals. Professionals here could include ECC officers and other such individuals.*

- Friends: (type number)
- Family: (type number)
- Professionals: (type number)

Appendix Figures & Tables

For Online Publication Only

Figures

Figure A2: Features Available in the App for Week 4

Figure A3: Enrollment of Participants - SMS Snapshot

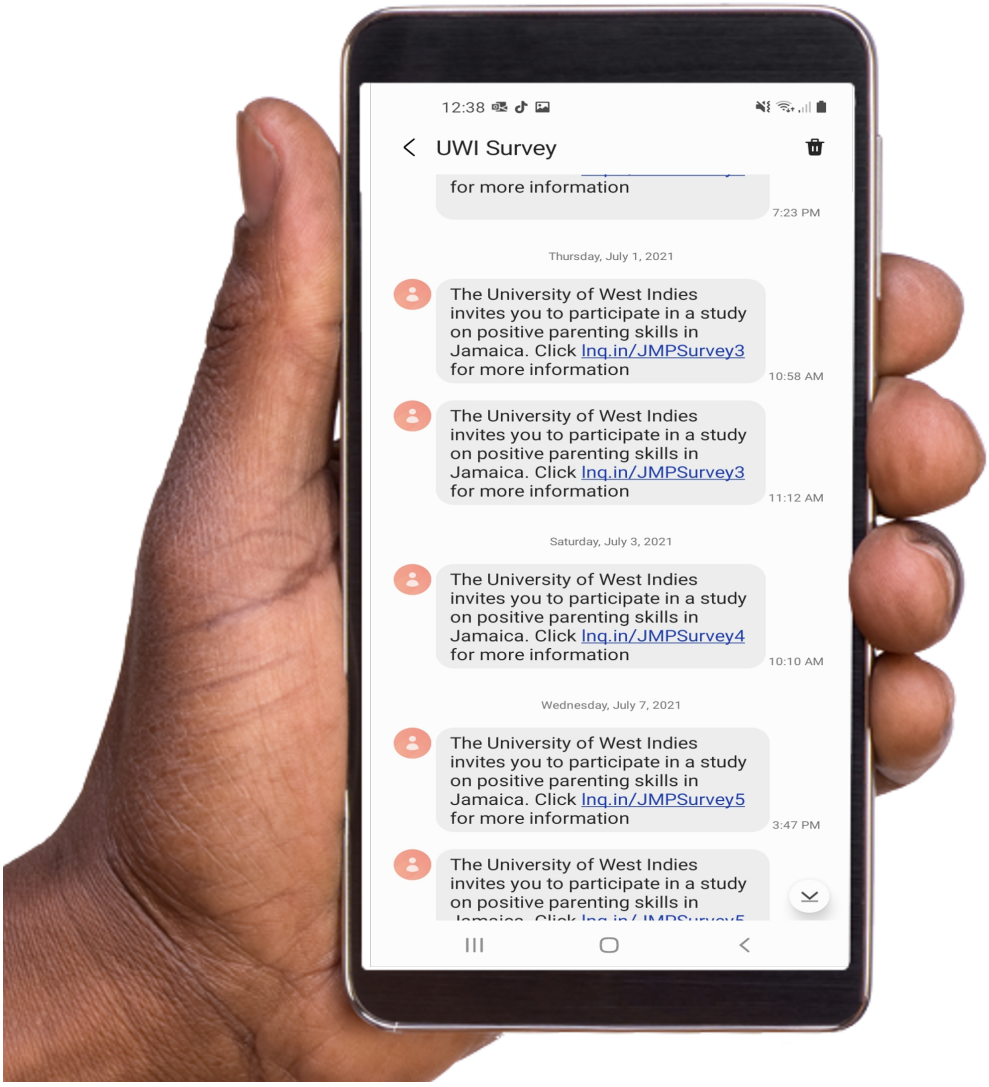


Figure A4: Enrollment of Participants

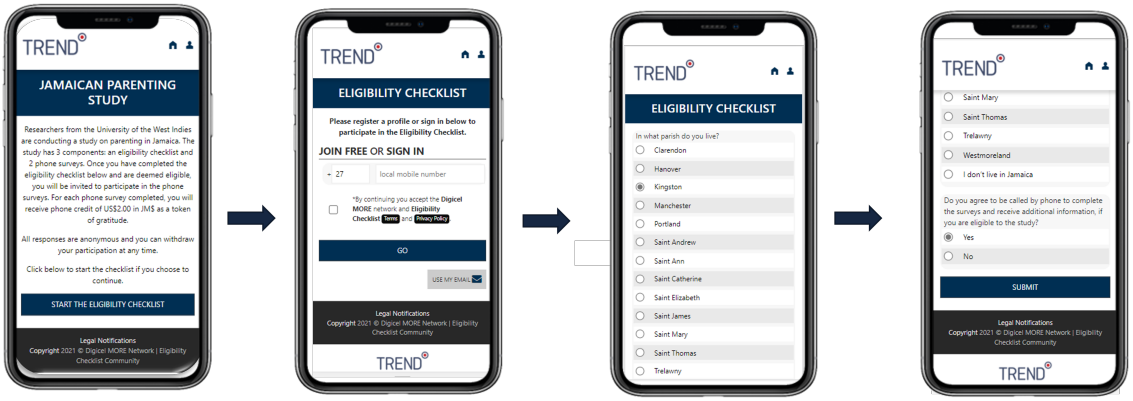


Figure A5: Enrollment Survey

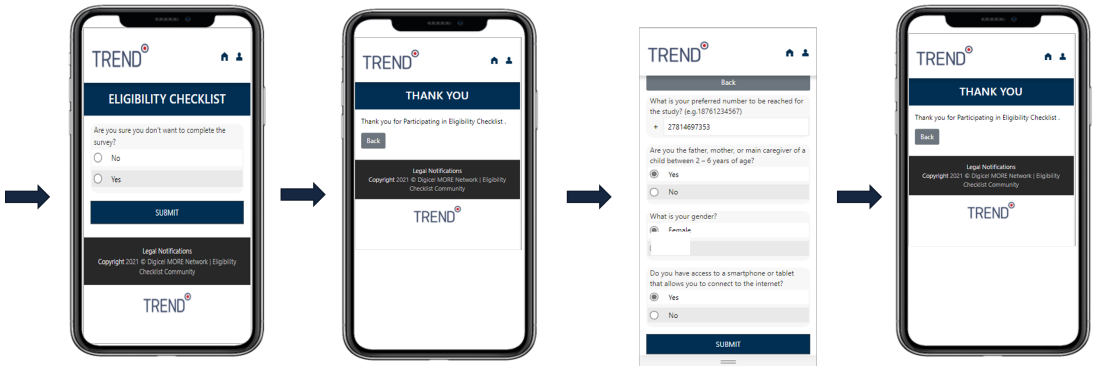


Figure A6: Distribution of Participants Compared to Population

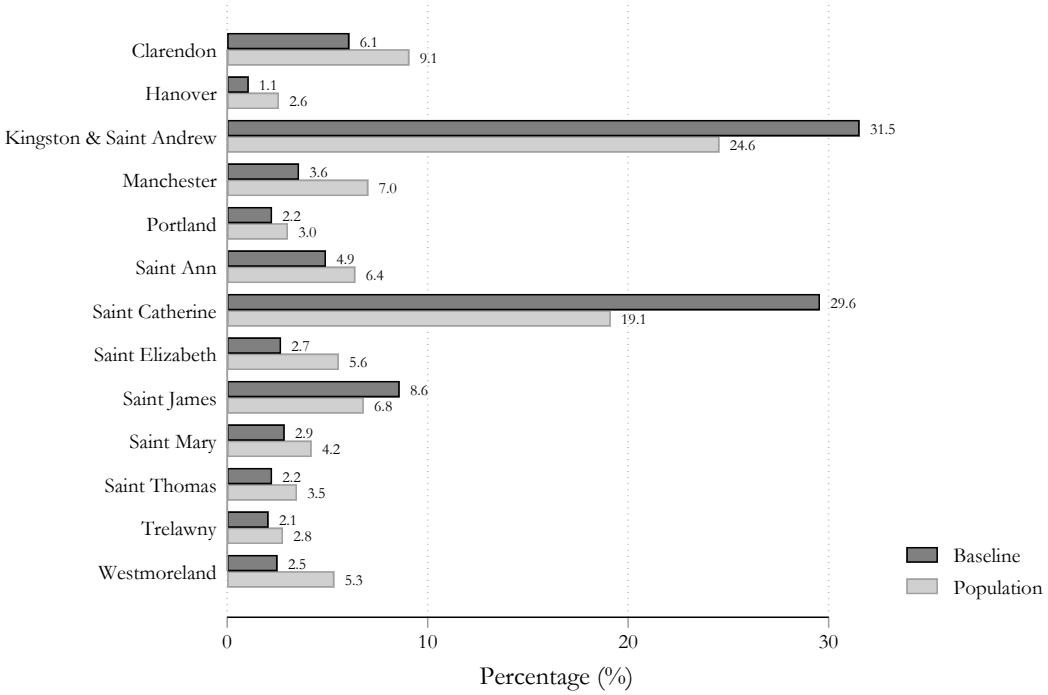
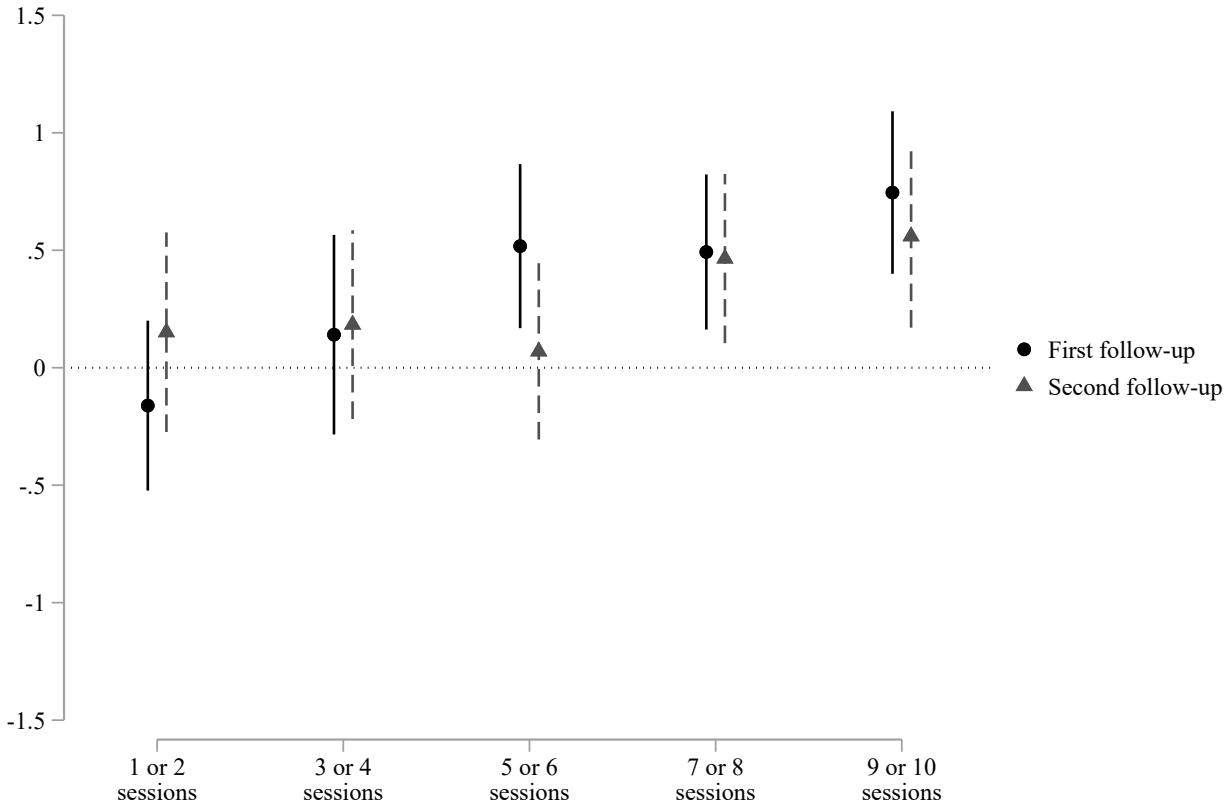
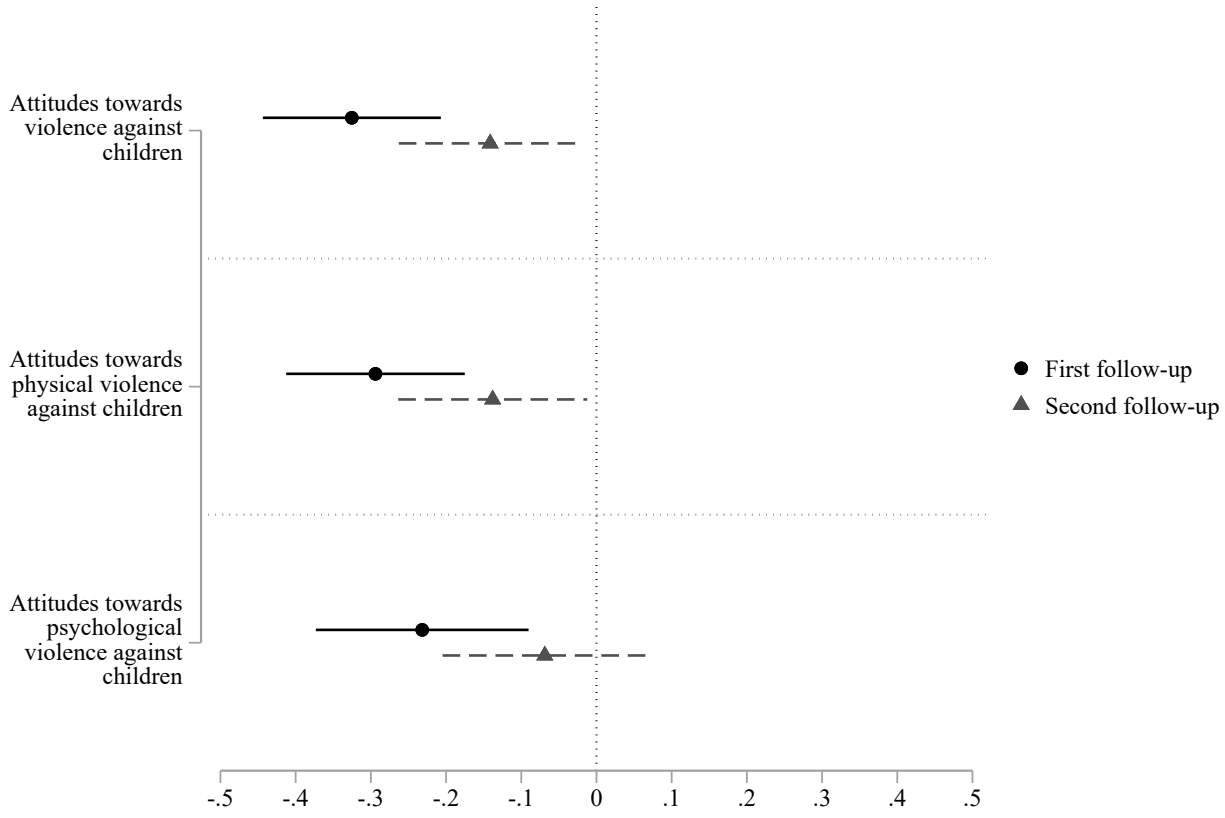


Figure A7: First Stage: Dose-response Regressions on Learning by Number of Sessions Attended



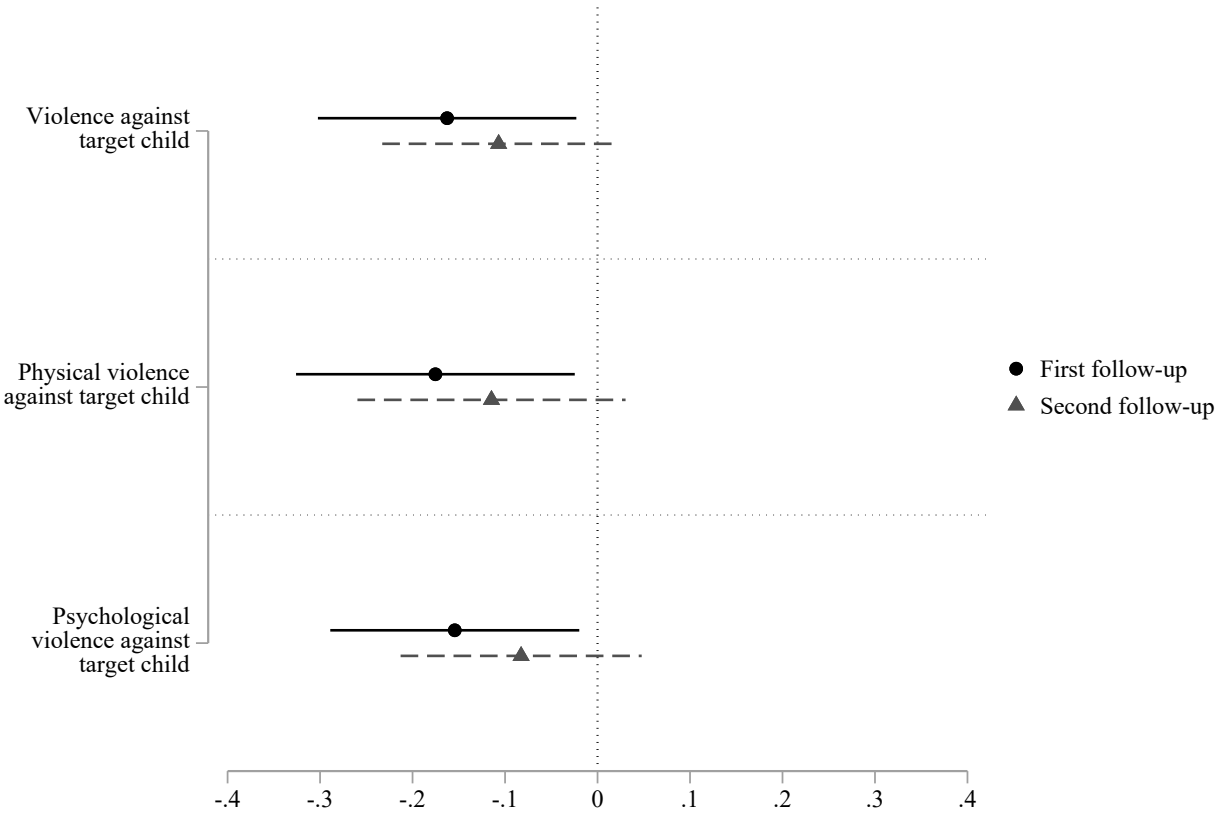
Notes: This figure plots point estimates and corresponding 95% confidence intervals from OLS regressions of the information index (dependent variable) on the number of sessions attended by caregivers (the independent variable). These dose-response regressions are only run with caregivers in the treatment group. To improve precision, we group the number of sessions attended into five categories (with attendance at zero sessions being the omitted base category). The black circles and corresponding solid lines represent the point estimates and 95% confidence intervals from the first follow-up survey. The gray triangles and corresponding dotted lines represent the point estimates and 95% confidence intervals from the second follow-up survey. Standard deviation units are used for the y-axis.

Figure A8: ITT Impacts on Caregiver Attitudes Related to Violence Towards Their Child (Balanced Panel)



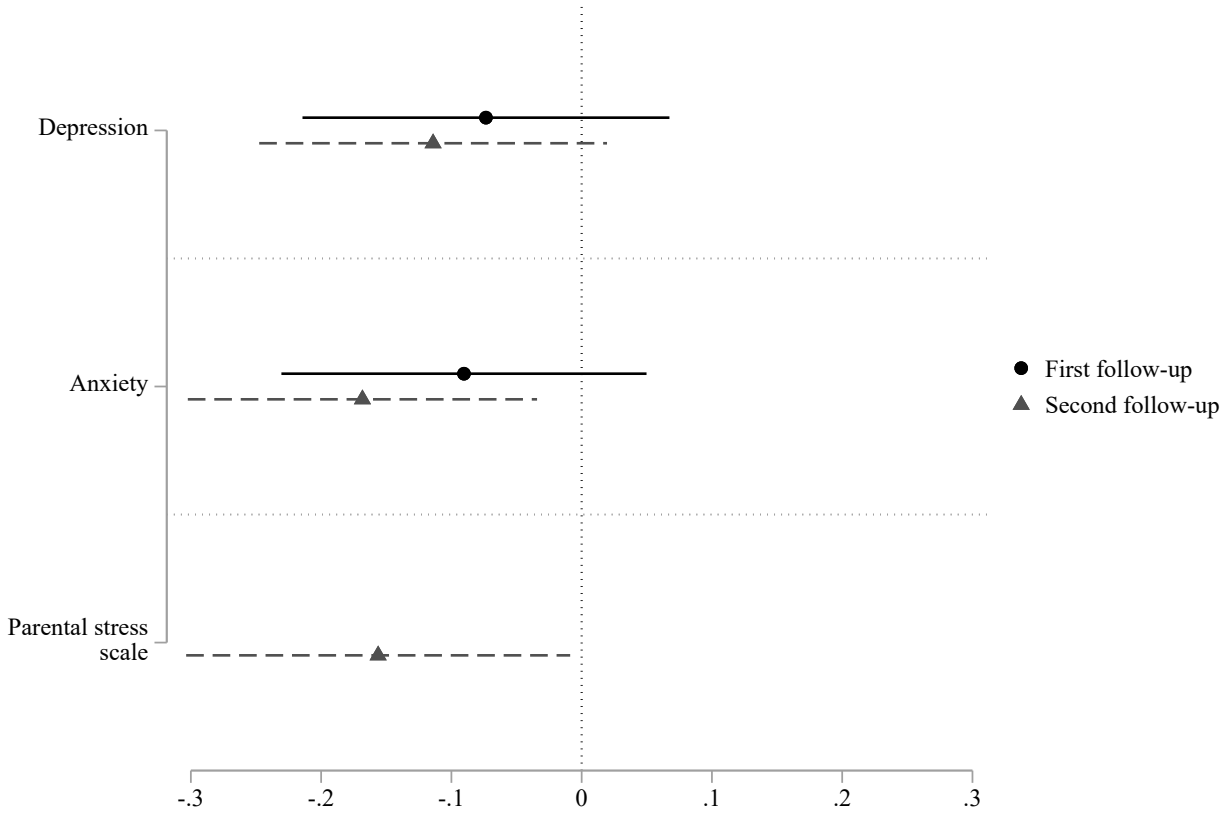
Notes: This figure presents OLS estimates of the short and medium-term intent-to-treat impacts on caregivers' attitudes relating to violence against children (indices). This figure uses a balanced panel of caregivers present at both first and second follow-ups. Each outcome consists of a standardized index estimated following Anderson (2008) and standardized relative to the control group. For a detailed description of the indices, see Section 3.3.2. The black circles and corresponding solid lines represent the point estimates and 95% confidence intervals from the first follow-up survey. The gray triangles and corresponding dashed lines represent the point estimates and 95% confidence intervals from the second follow-up survey. Standard deviation units are used for the x-axis.

Figure A9: ITT Impacts on Caregiver Behaviors Related to Violence Toward the Target Child (Balanced Panel)



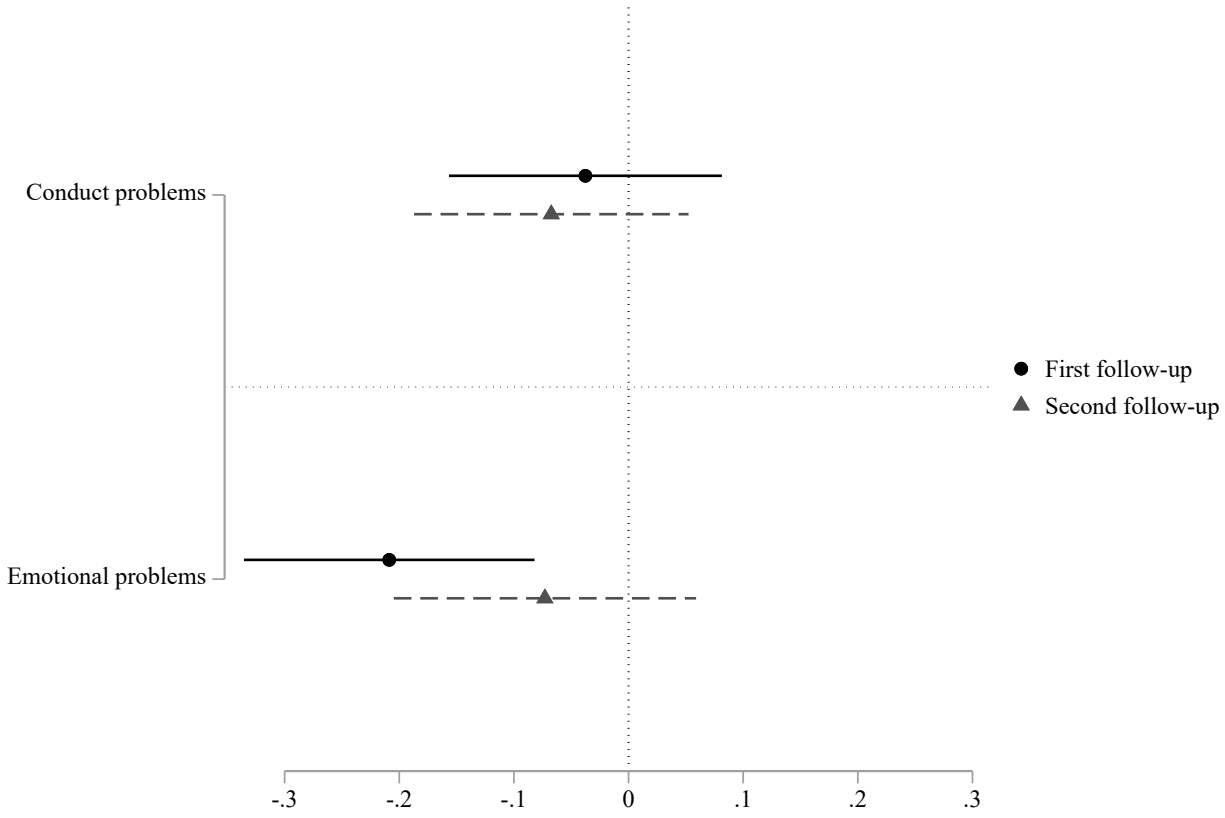
Notes: This figure presents OLS estimates of the short and medium-term intent-to-treat impacts on caregivers' behaviors (indices) relating to violence against the target child. This figure uses a balanced panel of caregivers present at both first and second follow-ups. Each outcome consists of a standardized index estimated following Anderson (2008) and standardized relative to the control group. For a detailed description of the indices, see Section 3.3.2. The black circles and corresponding solid lines represent the point estimates and 95% confidence intervals from the first follow-up survey. The gray triangles and corresponding dashed lines represent the point estimates and 95% confidence intervals from the second follow-up survey. Standard deviation units are used for the x-axis.

Figure A10: ITT Impacts on Caregivers' Depression, Anxiety, & Stress (Balanced panel)



Notes: This figure presents OLS estimates of the short and medium-term intent-to-treat impacts of the intervention on caregivers' depression (measured using PHQ-2 at both follow-ups), anxiety (measured using GAD-2 at both follow-ups), and parental stress (measured using PSS-18 only at second follow-up). This figure uses a balanced panel of caregivers present at both first and second follow-ups. Each outcome consists of a standardized index estimated following Anderson (2008) and standardized relative to the control group. For a detailed description, see Section 3.3.2. The black circles and corresponding solid lines represent the point estimates and 95% confidence intervals from the first follow-up survey. The gray triangles and corresponding dashed lines represent the point estimates and 95% confidence intervals from the second follow-up survey. Standard deviation units are used for the x-axis.

Figure A11: ITT Impacts on Child Conduct and Emotional Problems (Balanced Panel)



Notes: This figure presents OLS estimates of the short and medium-term intent-to-treat impacts of the intervention on the target child's conduct and emotional problems. This figure uses a balanced panel of caregivers present at both first and second follow-ups. Conduct and emotional problems were measured using caregivers' responses to the Strengths and Difficulties Questionnaire (SDQ) instrument. Each outcome consists of a standardized index estimated following Anderson (2008) and standardized relative to the control group. For a detailed description, see Section 3.3.2. The black circles and corresponding solid lines represent the point estimates and 95% confidence intervals from the first follow-up survey. The gray triangles and corresponding dashed lines represent the point estimates and 95% confidence intervals from the second follow-up survey. Standard deviation units are used for the x-axis.

Tables

Table A1: Structure of Virtual Parenting Programme

Session	TOPICS COVERED EACH SESSION
Session 1	PRAISING YOUR CHILD - Importance of praising your child - How to praise your child - What to praise your child for PRAISING YOURSELF: Importance of praising yourself for being a good parent
Session 2	INTRODUCING IRIE TIME - The importance of Irie Time - How to follow your child's lead in play - How to talk about what your child is doing - Using Respond, Describe, and Praise when playing with your child - Ideas for Irie Time activities
Session 3	GIVING YOUR CHILD POSITIVE ATTENTION THROUGHOUT THE DAY -Paying attention to positive behaviour during daily routines -Getting children involved in chores - Using Describe, Respond, and Praise during daily activities MODELLING: Modelling the behaviour you want / Being a good role model IRIE TIME: Playing with toys with your child
Session 4	GIVING CLEAR INSTRUCTIONS - How to give clear instructions (clear, specific, short, positively phrased, realistic, get child attention first) - Using labelled praise to praise child when they follow an instruction KNOW YOUR CHILD: Understand what your child likes/dislikes and factors that affect his/her behaviour IRIE TIME: looking at books with your child (using 'Going to School' book)
Session 5	TEACHING YOUR CHILD NEW SKILLS: - teach children how to follow the rules and expectations in the household. INDEPENDENCE: giving children independence CHOICES: -giving children choices. IRIE TIME: colouring with your child
Session 6	REASONS WHY CHILDREN MISBEHAVE: Identify why children misbehave ME-TIME: the importance of taking time to do something that you like to do. IRIE TIME: playing outside (ball, skipping, chasing) – how to do 'outside play activities' during Irie Time.
Session 7	MANAGING YOUR EMOTIONS -How we feel affects the way we behave. -Link back to the importance of trying to understand why children misbehave -How to recognize our own emotions. -How to calm down when feeling angry. HELPING CHILDREN UNDERSTAND THEIR OWN EMOTIONS -Labelling children's emotions IRIE TIME: Looking at books with your child (Use Emotions Book)
Session 8	HOW TO MANAGE YOUR CHILD'S MISBEHAVIOUR USING WITHDRAW ATTENTION AND REDIRECT -How to redirect children -How to withdraw attention from attention seeking behaviours. -How to use redirect and withdraw attention together. IRIE TIME: singing & dancing during Irie Time
Session 9	HOW TO MANAGE YOUR CHILD'S BEHAVIOUR USING CONSEQUENCES AND CHILLAX -How to use chillax -Giving children consequences -Problem solving IRIE TIME: pretend play (how to do pretend play activities during Irie Time)
Session 10	REVIEW OF THE YELLLOW AND GREEN BLOCKS OF THE IRIE TOWER I AM AN IRIE PARENT GOAL SETTING

Table A2: Structure of Virtual Parenting Programme

Week #	Date	Message
1	Sept 20 – 24	**Message 1** Children love to be praised. When you praise your child for doing something good, they will want to do it again and again. Learn more at: LINK HERE FOR SESSION 1
		Message 2 When praising, describe exactly what your child did, praise him/her and use your child's name. Add a clap or hug to make praise extra special. LINK HERE FOR SESSION 1
		Message 3 Irie Challenge for the week: Praise your child every day for all the good things they do. Praise yourself for being an Irie Parent. LINK HERE FOR SESSION 1
2	Sept 27 - Oct 1	**Message 1** Irie time is when we play and have fun with our child, doing what they want. Irie Time makes our child feel special and loved. LINK HERE FOR SESSION 2
		Message 2 During Irie Time, we can play with toys, look at books and play games. We let our child choose what they want to do and follow their lead. LINK HERE FOR SESSION 2
		Message 3 Irie Challenge for the week: Have Irie Time with your child for at least 10 minutes every day. Have fun! Great job for being an Irie Parent. LINK HERE FOR SESSION 2
3	Oct 4 - 8	**Message 1** Give your child positive attention and praise throughout the day. This will help your child learn to behave well and to learn new things. LINK HERE FOR SESSION 3
		Message 2 We are role models for our children. Children copy our behaviour. We need to speak and act in ways that we want our child to speak and act. LINK HERE FOR SESSION 3
		Message 3 Irie Challenge this week: Give your child positive attention throughout the day. Have Irie Time every day. Good job for being an Irie Parent. LINK HERE FOR SESSION 3
4	Oct 11 – 15	**Message 1** Try to give your child clear instructions and praise them whenever they follow your instruction. Praise encourages positive behaviour. LINK HERE FOR SESSION 4
		Message 2 As parents we know our child best. We know when they are most likely to misbehave. This can help us to prevent bad behaviour. LINK HERE FOR SESSION 4
		Message 3 Irie Challenge this week: Give your child clear instructions and praise them when they do what you say. Have Irie Time every day. Awesome Job. LINK HERE FOR SESSION 4
5	Oct 18 – 22	**Message 1** We can teach our child the little rules we have in our house and teach them important daily skills. This makes our life easier. LINK HERE FOR SESSION 5
		Message 2 We can help our child to behave well and learn well by giving them simple choices and allowing them some independence. This makes them feel good. LINK HERE FOR SESSION 5
		Message 3 Irie Challenge: Teach your child one skill this week. Give your child choices and some independence. Praise yourself for being an Irie Parent. LINK HERE FOR SESSION 5

6	Oct 25 - 29	**Message 1** To prevent our child from misbehaving, we need to understand the reason why our child is behaving in a certain way. LINK HERE FOR SESSION 6
		Message 2 As parents we are very busy. We need to take some time out for ourselves. In Me Time we do something that relaxes us and makes us feel happy. LINK HERE FOR SESSION 6
		Message 3 Irie Challenge: If your child misbehaves, try to understand why so you can prevent the behavior next time. Have some Me Time every day. LINK HERE FOR SESSION 6
7	Nov 1 – 5	**Message 1** When we feel angry with our child, we can stop, think and calm down before dealing with the situation. We can be a good role model for our child. LINK HERE FOR SESSION 7
		Message 2 We can help our child understand and manage their feelings by naming the emotion and describing why our child is feeling that particular emotion. LINK HERE FOR SESSION 7
		Message 3 Irie Challenge: Find ways to calm down when you are angry. Name your child's emotions and explain why they feel that way. You are an Irie Parent. LINK HERE FOR SESSION 7
8	Nov 8 - 12	**Message 1** Sometimes our child misbehaves when they are having fun, exploring & copying us. We can redirect our child's attention away from these behaviors. LINK HERE FOR SESSION 8
		Message 2 Children may try to get attention by crying for things, complaining, and nagging. We can withdraw attention from these behaviors. LINK HERE FOR SESSION 8
		Message 3 Irie Challenge: Use redirect and withdraw attention to deal with misbehavior. Praise your child when they behave well. Great job Irie Parent. LINK HERE FOR SESSION 8
9	Nov 15 - 19	**Message 1** We can give consequences for more serious misbehaviour. Consequences work best when they are short and not too harsh. We can also use Chillax. LINK HERE FOR SESSION 9
		Message 2 Chillax & consequences work best at managing our child's behaviour when we give them praise throughout the day for the good things they do. LINK HERE FOR SESSION 9
		Message 3 Irie Challenge: Give your child positive attention & praise for all the good things they do throughout the day. Awesome! You are an Irie Parent. LINK HERE FOR SESSION 9
10	Nov 22 - 26	**Message 1** Irie parents make Irie Homes for their children. Children feel safe, secure and loved and they have lots of opportunities to learn and play. LINK HERE FOR SESSION 10
		Message 2 Go back through the previous sessions. Set yourself parenting challenges each week to help you to continue making an Irie Home. Great job! LINK HERE FOR SESSION 10
		Message 3 CONGRATULATIONS. You have completed the Irie Homes Toolbox sessions. Praise yourself for being an Irie Parent and for making an Irie Home. LINK HERE FOR SESSION 10

Table A3: Survey Modules

Survey Modules	Baseline	First Follow up	Second Follow up
Caregivers Outcomes			
Attitudes to Violence Against Children	X	X	X
Violence (physical and psychological) against target child	X	X	X
Violence (physical and psychological) against eldest child	X	X	X
Depression, Sleep and Anxiety	X	X	X
Parental Stress Scale			X
Caregiver involvement in play and learning activities	X	X	X
Child Behavior			
Conduct and emotional problems (SDQ)	X	X	X
Mechanisms			
Brief Parental Self-Efficacy Scale (BPSES)		X	
Parental Self-Efficacy [From TOPSE – Discipline & Self-Acceptance]			X
Support networks	X	X	
Caregiver and target child socio-demographic characteristics			
Household Roster	X	X	X
Social Desirability Bias			X
Intervention take up and learning			
Information Module		X	X
Receipt of Parenting Support (+ take-up)		X	X

Table A4: Strata Composition

Strata	Frequency	Percentage
SMS - female	875	78.62
SMS - male	163	14.65
Principals or social media - female	70	6.29
Principals or social media - male	5	0.45
Total	1,113	100.00

Notes:This table shows the gender decomposition of the different recruitment sources for all participants in the intervention. The first two rows correspond to the participants recruited via SMS messages. Rows three and four correspond to participants recruited via school principals or social media. The columns show the absolute and relative frequencies associated with those variables.

Table A5: Descriptive Statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Panel A. Caregiver's characteristics					
Age	33.24	7.49	18.00	69.00	1,113
Gender - female (%)	0.85	0.35	0.00	1.00	1,113
Education level completed	14.33	2.75	6.00	19.00	1,113
Marital status - married (%)	0.37	0.48	0.00	1.00	782
Employed (%)	0.79	0.41	0.00	1.00	975
Income in the past month (USD)	882.16	1,114.53	0.00	6,424.67	877
Household total size	4.58	1.95	2.00	16.00	1,113
Total number of children 17 or below	1.91	1.03	1.00	8.00	1,113
Food insecurity - Eating less (%)	0.24	0.43	0.00	1.00	1,053
Panel B. Child's characteristics'					
Target child 2-6 years old - Age	4.12	1.43	2.00	6.00	1,113
Target child 2-6 years old - Female (%)	0.49	0.50	0.00	1.00	1,113
Panel C. Primary Outcomes					
Violence against target child (index)	-0.01	0.95	-0.98	5.71	1,112
Avg. # of days with discipline attitudes - target child	1.33	1.19	0.00	7.00	1,112
Violence against other child in the hh (index)	-0.01	0.99	-0.78	5.70	450
Avg. # of days with discipline attitudes - oldest child	1.05	1.25	0.00	7.00	450
Attit. to viol against children (index)	0.02	1.02	-0.81	4.04	1,109
(%) agreement with violent attitudes	18.25	21.72	0.00	100.00	1,109
Panel D. Secondary Outcomes					
Conduct problems (index)	-0.01	0.99	-1.32	3.58	1,113
(%) Conduct problems	41.80	27.00	0.00	100.00	1,113
Emotional problems (index)	-0.01	0.96	-0.89	4.41	1,113
(%) Emotional problems	25.88	24.50	0.00	120.00	1,113
Depression (index)	-0.00	0.97	-0.81	2.91	1,113
(%) High depression	19.05	39.29	0.00	100.00	1,113
Anxiety (index)	0.04	0.99	-0.71	3.56	1,113
(%) High anxiety	15.18	35.90	0.00	100.00	1,113
Caregiver involvement in play and learning activ.(index)	0.02	0.99	-3.19	1.35	1,104
Average number of good parenting practices	4.75	1.61	0.00	7.00	1,104
Panel E. Mechanisms					
Borrowing money support (index)	2.32	2.59	0.00	15.00	1,053
Parenting issues support (index)	2.65	2.74	0.00	20.00	1,068
Belong to parent support group (%)	0.16	0.37	0.00	1.00	1,113

Notes: This table shows descriptive statistics of the available variables at baseline for the sample of participant caregivers. Panels A and B present characteristics of the caregivers and their children. Panels C and D show the statistics for the primary and secondary outcomes, while Panel E shows the mechanisms measures. Columns Mean and Std. Dev describes the mean and the standard deviation associated with each of the variables in the first column. Min and Max describe the smallest and largest values associated with each variable, and column N describes the total number of caregivers in the treatment group to which the corresponding row variable applies.

Table A6: Comparison of Study Sample with Representative Survey

	JLCS 2019					WBG Survey					p-value
	Mean	Sd	Min.	Max.	N	Mean	Sd	Min.	Max.	N	
Panel A. Caregiver's characteristics											
Age	36.9	11.41	15	80	190087	33.24	7.49	18	69	1113	0.000
Gender - female (%)	0.89	0.31	0	1	190087	0.85	0.35	0	1	1113	0.076
Education	13.35	2.64	4	22	190087	14.33	2.75	6	19	1113	0.000
Marital status - married (%)	0.21	0.41	0	1	190087	0.37	0.48	0	1	782	0.000
Employed (%)	0.62	0.48	0	1	190087	0.79	0.41	0	1	975	0.000
Household total size	4.61	1.86	2	14	190087	4.58	1.95	2	16	1113	0.793
Total number of children 17 or below	2.23	1.19	1	7	190087	1.91	1.03	1	8	1113	0.000
Panel B. Child's characteristics											
Target child 2-6 year old - age	4.16	1.37	2	6	190629	4.12	1.43	2	6	1113	0.651
Target child 2-6 year old - female (%)	0.47	0.50	0	1	190629	0.49	0.50	0	1	1113	0.518

Notes: This table compares relevant descriptive statistics between our main sample and Jamaica Survey of Living Conditions restricted to caregivers with a child aged between two and six years. Panel A contains demographic variables associated with the caregiver, while Panel B shows information about the number of children between two and six years old under each caregiver and its decomposition by gender. The last column shows the p-value associated with the null hypothesis of the mean values across both groups being the same.

Table A7: Text Messages for Control Group

No.	Message
1	Thank you for participating in our parenting survey. You will now receive 3 SMS/week with tips on how to keep you and your child safe from COVID. Stay safe
2	COVID-19 Tip: Remind children to avoid sharing food, toys, pencils, books with friends.
3	COVID-19 Tip: Encourage children to wash their hands often with soap and water.
4	COVID-19 Tip: Remind children to avoid touching their face during COVID-19.
5	COVID-19 Tip: Wash your hands regularly when interacting with children
6	COVID-19 Tip: Avoid crowded spaces and close contact with others when traveling with children.
7	COVID-19 Tip: Keep children a safe distance from anyone with a cold or flu symptoms.
8	COVID-19 Tip: Remind children two years and over to avoid touching masks to reduce risks of contamination.
9	COVID-19 Tip: Remember to have children two years or older wear a mask outdoors and supervise mask wearing.
10	COVID-19 Tip: Show children 2 years and older, proper way to wear a mask and supervise them during mask wearing.
11	COVID-19 Tip: For children over two years, discard single use masks after each use and throw mask away without delay
12	COVID-19 Tip: Teach children proper ways to cover nose and mouth, when sneezing or coughing.
13	COVID-19 Tip: Remember to follow the COVID-19 protocols of the Ministries of Health and Education
14	COVID-19 Tip: Ensure children two years and older wear mask properly over nose, mouth, chin, it is secure, and they are supervised.
15	COVID-19 Tip: Remember to regularly wash reusable masks used by children over two years old.
16	COVID-19 Tip: Remove face masks of children in the right way, by removing from the ties and not touching the front.
17	COVID-19 Tip: For children over two years, remind them to wash their hands with soap and water after touching a used mask.
18	COVID-19 Tip: Ensure children have their own resources, to limit sharing with others
19	COVID-19 Tip: Remind children to avoid sharing food, toys, pencils, books with friends.
20	COVID-19 Tip: Dispose of single use face masks right after removal in a closed bin.
21	COVID-19 Tip: Remember to keep child at home or see a doctor if they are unwell with fever and cough.
22	COVID-19 Tip: Avoid crowded spaces and close contact with others when traveling with children.
23	COVID-19 Tip: Teach children to throw tissues used for sneezing or coughing into closed bin right after use, and wash hands.
24	COVID-19 Tip: Remember to teach your children over two years the proper way to wash their hands.
25	COVID-19 Tip: Remind children to keep safe distance from non-family members.
26	COVID-19 Tip: Remind children two years and over to wash their hands after coughing or sneezing.
27	COVID-19 Tip: Remember to clean your phones before giving to children to play.
28	COVID-19 Tip: Clean and disinfect high touch areas around the home used by children
29	COVID-19 Tip: Teach children how to properly wash hands with soap and water.
30	COVID-19 Tip: Regularly disinfect or wash toys and resources of children.
31	COVID-19 Tip: Wash your hands regularly when interacting with children.

Notes: This table enumerates the different COVID-19-related weekly SMS tips received by the caregivers in the treatment group.

Table A8: Extensive Margin ITT Impacts on Caregiver Attitudes and Behaviors (First Follow-up)

	Caregivers survey				
	Physical punishment needed	**Slap when misbehave**	**Beat to stop unwanted behavior**	**Shout and yell for obedience**	**Shout, yell, thretan to slap not harmful**
Panel A: Att. to VAC.					
Treatment	-0.039**	-0.062**	-0.051***	-0.023	-0.027
	(0.02)	(0.02)	(0.02)	(0.02)	(0.03)
Observations	916	847	882	886	870
R^2	0.11	0.23	0.13	0.11	0.12
Control Mean	0.121	0.226	0.097	0.077	0.329
	Hit with bare hand	**Hit with an object**	**Yelled**	**Threaten to send child away**	**Threaten to hit**
Panel B: Violence against target child					
Treatment	-0.097***	-0.026	-0.101***	0.040**	-0.055*
	(0.03)	(0.02)	(0.03)	(0.02)	(0.03)
Observations	916	885	934	884	928
R^2	0.15	0.03	0.21	0.10	0.12
Control Mean	0.395	0.074	0.745	0.090	0.675
	Hit with bare hand	**Hit with an object**	**Yelled**	**Threaten to send child away**	**Threaten to hit**
Panel C: Violence against all the children in HH					
Treatment	-0.087***	-0.024*	-0.088***	0.018	-0.082***
	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)
Observations	1226	1191	1252	1190	1246
R^2	0.14	0.03	0.20	0.09	0.13
Control Mean	0.365	0.080	0.729	0.102	0.671

Notes: This table shows the estimated short-term extensive margin impacts for each of the components in the caregivers' attitudes and behavior indices described in Figure 3. We present the estimated coefficient β_1 from the specification (1) in all the panels. "Treatment" is a dummy variable taking a value of one if the observed caregiver is in the treatment group, while "Control Mean" is the mean of the outcome for the control group. Sample size in each specification varies according to the number of observations available for each outcome. All results correspond to the first follow-up of the intervention. All dependent variables are measured as dummy variables whose descriptions are available in Appendix A2. Panel A describes the results of caregivers' attitudes toward VAC. The first three columns correspond to attitudes toward physical violence, while the last two columns correspond to psychological violence. These components are framed as Yes/No questions regarding the following statements: (1) In order to bring up, raise, or educate a child properly, the child needs to be physically punished. (2) good parent slaps their child when they misbehave. (3) When a child is beaten, he/she will stop doing the unwanted behavior. (4) Shouting and yelling make the child more obedient. (5) Shouting, yelling, and threatening to slap will harm the child. Panel B shows the extensive margin short-term impacts for the five components that comprise the violence against target child index, VATC. Panel C shows analog results to panel B for the case of Violence against all the children in the household. Panel B and C share the same components. These are framed as Agree/Disagree responses to the questions regarding the following statements: (1) hit the child on the bottom, hand, arm, or legs with their bare hands (2) hit the child with something like a belt, hairbrush, stick, or some other hard object, (3) shout, yell, or scream at their child, (4) threaten to send the child away in response to behavior perceived as inappropriate, (5) threaten to hit the child in response to behavior perceived as inappropriate. All specifications include strata-fixed effects for the four strata. Strata were defined as the cross between the gender of the caregiver (male or female) and the mode of recruitment into the study (SMS messages campaign or ECC/Principal referral and social media campaign). Heteroskedasticity-robust standard errors are reported in parenthesis. $^{*}\ p < 0.10$, $^{**}\ p < 0.05$, $^{***}\ p < 0.01$.

Table A9: Extensive Margin ITT Impacts on Caregiver Involvement in Play & Learning Activities (First Follow-up)

	Caregivers survey					
	Play	**Read books**	**Songs**	**Drawing**	**Praise**	**Index**
Treatment	0.010	0.003	0.018	-0.003	0.005	0.045
	(0.01)	(0.02)	(0.02)	(0.02)	(0.01)	(0.06)
Observations	934	924	919	912	929	945
R^2	0.05	0.04	0.10	0.04	0.02	0.08
Control Mean	0.943	0.877	0.863	0.897	0.976	-0.000

Notes: This table shows the estimated short-term extensive margin impacts for caregivers' play and learning activities. We present the estimated coefficient β_1 from the specification (1) in all the panels. "Treatment" is a dummy variable taking a value of one if the observed caregiver is in the treatment group, while "Control Mean" is the mean of the outcome for the control group. Sample size in each specification varies according to the number of observations available for each outcome. All results correspond to the first follow-up of the intervention. All dependent variables are measured as dummy variables whose descriptions are available in Appendix A2. The variables are Yes/No questions about activities performed during the past seven days: (1) played, (2) read books, (3) sang songs, (4) drew with or (5) praised their children. Column (6) shows results using a standardized index estimated following Anderson (2008) and standardized relative to the control group. All specifications include strata-fixed effects for the four strata. Strata were defined as the cross between the gender of the caregiver (male or female) and the mode of recruitment into the study (SMS messages campaign or ECC/Principal referral and social media campaign). Heteroskedasticity-robust standard errors are reported in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Attrition Analysis

	First follow-up		Second follow-up	
	(1) In endline β/se	(2) In endline β/se	(3) In endline β/se	(4) In endline β/se
Treatment	0.047**	-0.090	0.008	-0.273
	(0.02)	(0.18)	(0.03)	(0.24)
Age		0.002		0.008***
		(0.00)		(0.00)
Female		0.037		-0.119
		(0.10)		(0.12)
Education level completed		0.007		0.013*
		(0.01)		(0.01)
Marital status - married (%)		-0.019		-0.049
		(0.04)		(0.05)
Employed (%)		-0.041		-0.112**
		(0.03)		(0.05)
Income in the past month (USD)		-0.000		-0.000**
		(0.00)		(0.00)
Household total size		-0.007		0.015
		(0.01)		(0.01)
Total number of children 17 or below		0.021		0.006
		(0.02)		(0.03)
Food insecurity - Eating less (%)		0.032		-0.112**
		(0.04)		(0.05)
Viol. against target child (index)		-0.004		0.020
		(0.02)		(0.02)
Caregiver involv. in play and learning activ. (index)		-0.007		-0.003
		(0.01)		(0.02)
Depression (index)		0.006		0.027
		(0.02)		(0.02)
Anxiety (index)		-0.017		-0.016
		(0.02)		(0.02)
Conduct problems (index)		0.019		0.024
		(0.02)		(0.02)
Emotional problems (index)		-0.003		-0.020
		(0.02)		(0.02)
Att. to violence against children (index)		-0.001		0.004
		(0.02)		(0.02)
Treatment × Age		0.003		0.004
		(0.00)		(0.00)
Treatment × Female		-0.033		0.050
		(0.07)		(0.09)
Treatment × Education level completed		0.009		0.001
		(0.01)		(0.01)
Treatment × Marital status - married (%)		-0.024		-0.038
		(0.05)		(0.07)
Treatment × Employed (%)		0.021		0.057
		(0.05)		(0.07)
Treatment × Income in the past month (USD)		-0.000		0.000
		(0.00)		(0.00)
Treatment × Household total size		-0.009		-0.016
		(0.02)		(0.02)
Treatment × Total number of children 17 or below		0.004		0.039
		(0.03)		(0.04)
Treatment × Food insecurity - Eating less (%)		-0.070		0.102
		(0.05)		(0.07)
Treatment × Viol. against target child (index)		0.007		-0.027
		(0.02)		(0.03)
Treatment × Caregiver involv. in play and learning activ. (index)		0.020		0.011
		(0.02)		(0.03)
Treatment × Depression (index)		-0.025		-0.023
		(0.02)		(0.03)
Treatment × Anxiety (index)		-0.010		0.009
		(0.02)		(0.03)
Treatment × Conduct problems (index)		-0.016		-0.017
		(0.02)		(0.03)
Treatment × Emotional problems (index)		0.023		0.014
		(0.02)		(0.03)
Treatment × Att. to violence against children (index)		-0.001		0.015
		(0.02)		(0.03)
Constant	0.843***	0.684***	0.633***	0.377*
	(0.02)	(0.15)	(0.02)	(0.20)
Observations	1113	1113	1113	1113
R^2	0.01	0.04	0.00	0.06
Q test pvalue		0.897		0.916

Notes: This table shows the differences in attrition between treatment and control groups for each follow-up. The dependent variable "In Endline" in all columns is a dummy indicating if a caregiver responded to the follow-up surveys. Models 1 and 3 measure the impact of the treatment on the follow-up survey respondent. Models 2 and 4 measure the impact of any demographic characteristics or outcome variables measured at the baseline on the probability of completing the follow-ups. All regressions include strata fixed-effects. Strata were defined as the cross between the gender of the caregiver (male or female) and the mode of recruitment into the study (SMS messages campaign or ECC/Principal referral and social media campaign). For the first follow-up we imputed the mean for the observations with missing values (1099 with complete observation, only 14 cases). Heteroskedasticity-robust standard errors are reported in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Robustness Check: Lee Bounds for Attrition Analysis (First Follow-up)

	Primary hypotheses			
	Attitudes to violence against children (index) β/se	Violence against target child (index) β/se	Violence against children in the household (index) β/se	Caregiver involvement in play and learning activ. (index) β/se
Treatment				
lower	-0.301***	-0.291***	-0.202***	-0.049
	(0.07)	(0.08)	(0.07)	(0.08)
upper	-0.168***	-0.110	-0.161**	0.141
	(0.06)	(0.08)	(0.06)	(0.09)
Observations	1111	1095	1273	1105

Notes: This table shows the Lee bounds associated with the estimates for treatment effects. Lower and Upper bounds are provided for each of the four indices describing the main outcomes of the intervention explained in Section 3.3.2. P-values for non-significant bounds: (i) Violence against target child (index) upper bound: 0.177; (ii) Caregiver involvement in play and learning activ. (index) lower bound: 0.557; upper bound: 0.108. Sample size in each specification varies according to the number of observations available for each outcome. Heteroskedasticity-robust standard errors are reported in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Double LASSO for Selection of Controls

	Primary hypotheses			
	Attitudes to violence against children (index)	Violence against target child (index)	Violence against children in the household (index)	Caregiver involvement in play and learning activ. (index)
Panel A: First follow-up				
Treatment	-0.186***	-0.112**	-0.132***	0.017
	(0.05)	(0.06)	(0.05)	(0.06)
Observations	977	943	1265	945
# of controls selected	4	5	6	3
Control mean	-0.000	-0.000	0.003	0.000
Panel B: Second follow-up				
Treatment	-0.142**	-0.107*	-0.125**	0.011
	(0.06)	(0.06)	(0.05)	(0.07)
Observations	696	681	914	676
# of controls selected	3	4	5	3
Control mean	-0.000	0.000	-0.011	-0.000

Notes: This table shows the estimated treatment effects associated with each of the main outcome variables described in Section 3.3.2. Each specification includes the Double Lasso suggested controls for each of the main outcomes. Panel A describes the short-term effects corresponding to the first follow-up, while Panel B presents the analog results for the second follow-up. "Number of Controls Selected" refers to the total number of controls included in the specification. "Control Mean" describes the mean of the outcome for the control group. The sample size in each specification varies according to the number of observations available for each outcome. Heteroskedasticity-robust standard errors are reported in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A13: Relationship between Treatment Assignment & Social Desirability Bias (Second Follow-up)

	Social desirability module						
	Hard to go to work	Resentful if not my way	Few ocassions give up	Rebelling knowing I was wrong	Good listener always	Took advantage of someone	Willing to admit mistakes
Panel A: SDB Components 1-7							
Treatment	-0.044	0.079**	0.035	-0.048	-0.041	0.049	0.012
	(0.04)	(0.03)	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Observations	693	689	691	695	692	693	695
R^2	0.02	0.03	0.01	0.01	0.04	0.03	0.01
Control mean	0.65	0.73	0.63	0.29	0.89	0.71	0.86
	Get even rather than forget	Always courteous	Never irritated	Jealous of others fortune	Irritated by people	Never said something hurtful	SDB (index)
Panel B: SDB Components 8-13 + SDB Index							
Treatment	0.060*	0.023	-0.018	-0.003	0.007	-0.053	-0.016
	(0.03)	(0.03)	(0.04)	(0.03)	(0.04)	(0.04)	(0.08)
Observations	688	691	687	696	689	692	700
R^2	0.03	0.00	0.01	0.02	0.00	0.02	0.01
Control mean	0.700	0.843	0.476	0.836	0.605	0.484	0.000

Notes: This table shows the relationship between treatment assignment and each component of the Marlowe-Crowne Social Desirability Scale as described in Section 3.3.2. Each of the columns corresponds to a different dummy Marlow-Crowne variable capturing the respondent's approval of the statement. The last column of Panel B corresponds to a specification in which the Social Desirability Index is the dependent variable. The term "Control Mean" describes the mean of the outcome for the control group as measured at second follow-up. The sample size in each specification varies according to the number of observations available for each outcome. Heteroskedasticity-robust standard errors are reported in parenthesis. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: Comparison of Treatment Impacts Between First & Second Follow-ups

Variable	First follow-up β_{f1}	First follow-up SE_{f1}	Second follow-up β_{f2}	Second follow-up SE_{f2}	p-value $\beta_{f1} - \beta_{f2}$
Panel A. Primary outcomes					
Att. to viol. against children	-0.198	0.050	-0.144	0.061	0.497
Viol. against target child	-0.121	0.056	-0.127	0.062	0.950
Viol. against children in the hh.	-0.145	0.048	-0.149	0.053	0.959
Care. involv. in play and learn. activ.	0.019	0.057	0.006	0.067	0.885
Panel B. Primary outcomes (subindices)					
Att. to phys. viol. against children	-0.191	0.051	-0.150	0.063	0.615
Att. to psyc. viol. against children	-0.120	0.058	-0.075	0.068	0.612
Phys. viol. against target child	-0.136	0.060	-0.124	0.069	0.897
Psyc. viol. against target child	-0.101	0.055	-0.105	0.065	0.961
Panel C. Secondary outcomes					
Conduct problems	-0.030	0.049	-0.053	0.059	0.760
Emotional problems	-0.166	0.053	-0.051	0.065	0.170
Panel D. Effective learning					
Praising children helps	0.189	0.045	0.104	0.049	0.203
Imp. for parents to play w child	0.059	0.038	0.039	0.044	0.729
Clear instructions help	0.224	0.046	0.248	0.060	0.753
Understand why child misbehaves	0.063	0.037	-0.058	0.038	0.023
Calm down before disciplining	-0.011	0.039	-0.026	0.040	0.788
Withdraw att. from childs whining	0.624	0.068	0.480	0.079	0.168
Redirect rather than reprimand	0.278	0.053	0.165	0.058	0.150
Conseq. and timeout appropriate	0.080	0.040	0.071	0.044	0.870
Information module (index)	0.525	0.070	0.394	0.080	0.217
Panel E. Mechanisms					
Depression	-0.033	0.059	-0.124	0.067	0.309
Anxiety	-0.070	0.059	-0.157	0.068	0.330

Notes: This table presents a comparison of treatment impacts between the first and second follow-ups as estimated using our main specification (1). β refers to the estimated coefficients and SE refers to the estimated heteroskedasticity-robust standard errors. The last column presents the p-value for a test of difference in means between the estimated coefficients for the first and second follow-up.