Predictors of Attendance and the Impact of Attendance on Outcomes for a Parenting Programme in Two Southeast Asian Countries

Roselinde Katharina Janowski

Supervised by Professor Catherine L. Ward (University of Cape Town) &

Dr Jamie M. Lachman (University of Oxford)

Department of Psychology

University of Cape Town

August 2020

A minor dissertation submitted in partial fulfilment of the requirements for the award of the degree of Master of Arts in Psychological Research

COMPULSORY DECLARATION

This work has not been previously submitted in whole, or in part, for the award of any degree. It is my own work. Each significant contribution to, and quotation in, this dissertation from the work, or works, of other people has been attributed, and has been cited and referenced.

Signature: Signed by candidate Date: 7 August 2020

Abstract

***Background*:** Children living in low- and middle-income countries (LMICs) experience alarmingly high rates of maltreatment, frequently at the hands of caregivers. Group-based parenting programmes show promise for reducing and preventing child maltreatment, as well as for improving positive parenting, child behaviour problems, and caregiver mental health. However, parenting programmes can only benefit families if caregivers participate in them. Using secondary data, this study thus aimed to 1) identify factors that affect attendance and 2) investigate the impact of attendance on outcomes within two randomised controlled trials of Parenting for Lifelong Health (PLH) for Young Children for caregivers of children aged 2-9 years in Thailand (N = 120) and 2-6 years in the Philippines (N = 120). The interventions were delivered within existing service delivery systems in both countries, over eight weekly sessions (Thailand) or 12 sessions every second week (Philippines). ***Method*:** To address the first aim of this study, multivariable logistic regression models with robust sandwich estimators were used to examine family baseline characteristic as predictors of caregiver attendance in sessions. An exploratory approach was taken to test a range of factors that have previously been linked to attendance in parenting programmes, including economic and educational, social and health, parenting and child behaviour, and sociodemographic characteristics. To address the second aim, caregiver self-reports and observational assessments (Thailand only) from baseline, post-test, and follow-up were analysed using complier average causal effect (CACE) analyses to test the impact of attendance variability on the primary outcomes of child maltreatment, as well as secondary outcomes of positive parenting, dysfunctional parenting, child behaviour problems, and caregiver mental health. ***Results*:** Caregivers in Thailand attended 82.3% of sessions while those in the Philippines attended 61.8%. Overall, few baseline factors were significantly associated with attendance. In Thailand, caregivers who were less educated and those who were older were significantly

more likely to attend sessions. In the Philippines, caregivers who were less healthy, those that who used more emotional abuse, and those who had boys rather than girls were significantly more likely to attend. Notably, caregivers who experienced higher rates of intimate partner violence significantly attended 8% fewer sessions in the Philippines. A comparison of CACE estimates to intention-to-treat estimates at post-test and at follow-up showed greater benefits of the intervention amongst caregivers who attended more sessions. Specifically, the strongest intervention effects were found for caregivers who attended at least 75% of the programme. ***Conclusion****:* This study showed no evidence that disadvantages related to lower socio-economic status were associated with attendance, suggesting that it is possible for vulnerable families in LMICs to attend parenting programmes. However, developing retention strategies that target subgroups who are at greater risk of missing sessions is especially important as higher attendance at sessions is positively related to greater improvements in caregiver and child outcomes.

Keywords: Parenting programmes; child maltreatment; attendance; complier average casual effects, Southeast Asia.

Acknowledgements

I am thankful to so many people who made this thesis possible. First, I would like to extend my gratitude to the hundreds of families who participated in the RCTs, the facilitators who delivered the programmes, and the research assistants who spent many long hours collecting, capturing and cleaning the data that forms the basis of this work.

I would also like to thank the PLH research team involved in designing and implementing the RCTs. Thank you for giving me the opportunity to be part of this exciting research, even though I only joined recently. My initial thesis topic involved examining the barriers and facilitators of scaling up PLH in South Africa. However, due the COVID-19 pandemic and South Africa's extended lockdown, programme delivery and data collection became impossible. A special thank you goes to Lucie Cluver and Jamie Lachman for brainstorming alternatives and for helping me find a new topic which I am just as passionate about.

I would also like to thank everyone who provided assistance with the statistical analyses. Professor Colin Tredoux from the University of Cape Town who went out of his way to research CACE analysis even though he had never used it before. Professor Gerardo Meléndez-Torres from the University of Exeter for his advice on CACE analysis for longitudinal data and for methods of classifying noncompliers in the control group. Lastly, thank you to Yulia Shenderovich for the helpful list of resources, and for pointing me in the direction of an excellent blog on performing CACE analysis in R.

I would also like to extend my gratitude to the David and Elaine Potter Foundation both for their financial support throughout my master's and for their belief in my research and the research concerning preventing violence against children more broadly.

I would like to thank my supervisors for their continuous guidance and support. Cathy, thank you for standing by me throughout my honour's and now my master's, for

pushing me to do my best, and for providing kindness and advice whenever I needed it. Without you, I would not be where I am today. Jamie, thank you for stepping in at the last-minute and for offering so much of your invaluable time and expertise. I am eternally grateful for all the opportunities you have both given me to learn and grow as a researcher.

I am also grateful to my family and friends for their support and patience throughout this work. Thank you to Brittany McCreath for checking in on me and for always being there when I needed her. Thank you to Miguel Pereira for being my biggest supporter. I would not have been able to accomplish any of this without you by my side.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

Violence against children - encompassing all forms of physical, emotional, and sexual abuse, exploitation, maltreatment, and neglect (World Health Organization, 2014) - is a serious global problem, cutting across all levels of education, income, and culture (United Nations General Assembly, 2006). Findings using population-based data from 96 countries indicate that over one billion children aged 2-17 years are exposed to such violence every year (Hillis et al., 2016). However, a disproportionate burden of violence is experienced by children in low- and middle-income countries (LMICs), with incidence rates estimated to be highest in Asia at 64%, compared to 50% in Africa and 34% in Latin America (Hillis et al., 2016).

Exposure to violence in childhood has immediate and long-term consequences on a range of social, health, and economic outcomes including mental and physical health conditions, life expectancy, employment, crime, and intergenerational violence, representing an immense cost to society (Anda et al., 2006; Fang et al., 2012; Fang et al., 2015; Fang et al., 2017; Fry et al., 2018; Hughes et al., 2017; Knox et al., 2011). In response to the prevalence rate of violence against children and associated costs and consequences thereof, ending all violence against children has been identified as a core target for development within the United Nations 2030 Sustainable Development Goals (WHO, 2018). There is particular urgency to support the achievement of these goals in LMICs, such as the Southeast Asia region, where rates of violence against children are highest (Hillis et al., 2016; UNICEF East Asia and Pacific Regional Office, 2014), but social services to prevent child maltreatment are scarce (McCoy et al., 2020). Implementing effective and accessible prevention strategies that reach children and families affected by violence in these settings is therefore paramount.

Responsive, warm, and consistent care from parents and other caregivers are important protective resources in promoting children's health and development, and are especially essential in low-income and high-stress settings (Betancourt & Khan, 2008; Murphy et al., 2017; Tol et al., 2013).[1] Therefore, there is substantial need to promote effective parenting strategies and advance knowledge around parenting in LMIC settings. One promising approach to support and promote parenting is through social learning, theory-based parenting interventions. These interventions typically involve parents and caregivers actively learning new skills to enhance caregiver-child relationships through positive parenting as well as manage child behaviour problems through nonviolent discipline strategies (Leijten et al., 2018). Robust evidence suggests that these parenting programmes are effective at reducing and preventing violence against children (e.g. Barlow & Coren, 2018; Barlow et al., 2006; Chen & Chan, 2016; Mikton & Buchart, 2009). They have also been shown to contribute to reducing caregiver mental health problems, child behaviour problems, and adolescent risk behaviours (e.g. Furlong et al., 2013; Piquero et al., 2009). While most studies of parenting programmes have been conducted in high-income, Western settings, a growing body of evidence from LMICs and non-Western contexts also indicates promising effects (Knerr et al., 2013; McCoy et al., 2020). In light of this evidence, and as part of the efforts to achieve the 2030 Sustainable Development Goals, the WHO-led INSPIRE framework recommends parenting programmes as an important strategy for reducing and preventing violence against children (WHO, 2016).

Despite the promise of these programmes, expensive licencing fees and programme costs inhibit their widespread implementation in many LMICs (Mikton, 2012). In response to

[1] Parents and caregivers are used interchangeably in this thesis to refer to any person responsible for the care of a child regardless of biological relation.

the need for affordable and freely available programming, the Parenting for Lifelong Health (PLH) initiative was formed[2]. This initiative is a partnership between the WHO, UNICEF, academic institutions, and implementing agencies from the Global South. It is aimed at developing, rigorously testing, and disseminating a suite of violence prevention parenting programmes targeted to specific age groups across the child development spectrum. The programmes were developed with participatory input from families in the Global South for delivery by trained para-professional staff with supervisory support, have minimal equipment requirements, and are freely available under Creative Commons Licenses that prohibit any profit or commercial interests.

One of these programmes is PLH for Young Children, originally developed and tested as a 12-session, group-based parenting intervention for caregivers of children aged 2-9 years (Lachman et al., 2017; Ward et al., 2020). Core functional components focus on building positive caregiver-child relationships, positive reinforcement of positive behaviour, setting limits, non-violent discipline, and stress reduction (Lachman et al., 2016). A recent randomised controlled trial (RCT) with 296 caregivers in very low-income areas of Cape Town, South Africa, showed that the programme increased rates of positive parenting and reduced the risk of harsh parenting by caregivers (Ward et al., 2020). It has since been adapted and tested in other LMICs including Thailand and the Philippines – the contexts in which the current study takes place.

Thailand has made remarkable social and economic progress over the last four decades, transitioning from a low-income to an upper-middle-income country in 2011 (The World Bank, 2020). Nevertheless, like most other LMICs, prevalence rates of violence against children by parents and caregivers are estimated to be high. A nationally

[2] PLH website: http://www.who.int/violence_injury_prevention/violence/child/plh/en/.

representative survey of 28,653 households indicated that 75% of adults reported psychological and physical punishment of children aged 1-14 years during the last month, with violent discipline being most common amongst 3-4 year olds at 81% (National Statistical Office & United Nations Children's Fund, 2016). Further, a longitudinal study with parents of 7-10-year olds in the city of Chiang Mai found that 58% of girls and 72% of boys had experienced physical punishment over the past month (Lansford et al., 2014). Similarly, a study of 413 sixth-grade students in Bangkok showed that 64% reported experiencing verbal aggression from their caregivers, with 11% reporting verbal aggression on a daily basis (Isaranurug et al., 2001). To address these high prevalence rates of violence against children by parents and caregivers, the Universities of Oxford and Cape Town, in partnership with the Thai Ministry of Public Health (MOPH), UNICEF Thailand, and Clowns Without Borders South Africa, adapted and tested PLH for Young Children for delivery by nurses, mental health officers, child protection officers, and community health workers within the public health system in local community clinics in Udon Thani Province, Thailand.

The Philippines is a lower-middle-income country where prevalence rates of violence against children are also high. In an international study of six countries, Runyan and colleagues (2010) found that rates of physical abuse among the Philippine community was 76%. Similarly, a longitudinal study with parents of 7-10-year olds in Metro Manila showed that 71% of girls and 77% of boys had experienced some form of corporal punishment over the past month (Lansford et al., 2014). A recent nationally representative survey of 3,866 adolescents aged 13-24 years in the Philippines found that rates of violence against children among females were 78.4% and 81.5% among males (UNICEF, 2016). This study included physical abuse, emotional abuse, and neglect in the definition of “violence against children”, and showed that 66.3% of respondents experienced physical abuse during childhood, with 60% of these cases occurring within the home. Emotional abuse and neglect were also shown

to be high, with an estimated lifetime prevalence of psychological violence during childhood at 59.2% (UNICEF, 2016). In response to these high prevalence rates of child maltreatment, PLH for Young Children was adapted for families with children aged 2-6 years living in Metro Manila, Philippines, who are enrolled in a conditional cash transfer system run by the Philippine Department of Social Welfare and Development (Alampay et al., 2018). This conditional cash transfer is a social protection programme that provides cash grants to low-income families who participate in a range of health, educational and family development services.

The PLH for Young Children programme has recently been tested in RCTs in Thailand and the Philippines to examine its effectiveness of reducing violence against children and associated risk factors (Gardner et al., forthcoming; Lachman et al., forthcoming). While it is critical to establish the effectiveness of the adapted versions of PLH for Young Children, it is equally important to examine the role of participation, or "engagement" in these settings. Engagement in parenting programmes has been defined in various ways in the literature, including participant attendance at sessions, quality of participation during sessions, and level of home practice completion (Dumas et al., 2007). Of these, attendance is a fundamental and necessary antecedent to other indicators such as quality of participation in sessions or use of skills outside of the session (Berkel et al., 2018). Parenting programmes teach specific skills during sessions and caregivers have opportunities to observe, learn and practise these skills. Missing sessions not only results in missing specific programme content but also means missing opportunities to consolidate and practise new skills. Session attendance may therefore be critical for achieving programme success (Whittaker & Cowley, 2012).

Although studies in high-income countries (HICs; Patterson et al., 2002) and LMICs (Ward et al., 2015) have indicated that caregivers are interested in attending parenting

programmes, poor attendance is common. A recent review of 262 studies of parenting interventions in HICs showed that at least 51% of participants were lost to attrition, i.e., they dropped out of the programme before the last session (Chacko et al., 2016). This review also found that while the average attendance rate reported was 72%, some studies reported rates as low as 29%. Moreover, high rates of attrition and poor attendance have been reported even when participants receive childcare, transportation, or financial incentives (e.g. Baker et al., 2011; Duppong-Hurley et al., 2016).

Poor attendance and attrition are problematic for several reasons. Many parenting programmes, including PLH for Young Children, use group-based formats as their primary mode of delivery. Smaller groups can be less cost-effective (Axford et al., 2012) and might thus not be feasible for implementation in resource-constrained settings. Small groups can also render group activities and group discussions difficult, thereby limiting the opportunities participants have to actively engage in sessions (Laxman et al., 2019). Therefore, attrition and poor attendance have significant implications for widespread implementation as they diminish impact, consume scarce resources, and reduce reach and suitability of interventions (Spoth et al., 2002).

Given the significant implications of low or no attendance, it is critical to examine issues of caregiver attendance in parenting programmes before their wide-scale implementation. First, understanding how attendance is associated with intervention outcomes is useful as it may identify the "dosage" required to ensure positive programme effects or explain weak or null effects (Flay et al., 2005). Second, understanding factors associated with attendance, including barriers and resources to programme access, may help improve the feasibility of the interventions within the target population. That is, it can provide insight into which subgroups of caregivers are most likely to attend, and which caregivers are at risk of attrition. Together, this information can inform effective retention

strategies to increase attendance and ensure that as many families as possible benefit from programmes (Chacko et al., 2016).

Several studies have therefore investigated the issue of poor attendance in parenting programmes, seeking to identify factors that predict attendance (e.g. Baker et al., 2011; Dumas et al., 2007), and to understand how attendance is related to programme outcomes (e.g. Baydar et al., 2003; Nix et al., 2009). These studies have predominantly been conducted in HICs, and there is no known literature on attendance in parenting programmes in Southeast Asian settings such as Thailand or the Philippines. However, a recent study of PLH for Young Children (N = 296) investigated the factors associated with variations in attendance and the impacts of attendance on outcomes in South Africa (Wessels, 2017). This study found that while there were no social, contextual, behavioural, or demographic factors strongly associated with attendance, higher attendance at sessions was significantly related to more positive parenting at the end of the programme, but not with reductions in observed negative parenting behaviour (see below for more details). This study provides key insights into the potential resources and barriers to attendance in PLH for Young Children as well as the ways in which attendance influences programme outcomes. Nonetheless, there are key differences between the context of delivery in South Africa and that of Thailand and the Philippines. In South Africa, the programme was delivered by lay community workers in urban settings, while PLH for Young Children Thailand was delivered by health workers within the public health system in predominantly rural settings and PLH for Young Children Philippines was delivered by a mixture of professionals and paraprofessionals within a conditional cash transfer system in urban settings. Therefore, in addition to different social, cultural, and economic settings, implementation settings and delivery channels were vastly different, and it is thus unclear whether findings from South Africa will maintain across these new settings.

To build upon the work of Wessels (2017) and extend knowledge and understanding of attendance in parenting programmes in LMICs, this current study sought to a) advance evidence of the factors associated with attendance of parenting programmes in LMICs in concurrent studies of PLH for Young Children in Thailand and the Philippines, and b) investigate in what ways attendance was associated with change in PLH for Young Children within the different delivery contexts in Thailand and the Philippines. To inform study hypotheses, the literature review that follows will first synthesise research on the possible predictors of attendance in parenting programmes, and then examine the potential impact of session attendance on targeted outcomes.

Correlates of Attendance in Parenting Programmes

Given the lack of research on factors associated with greater/lower attendance within the context of parenting programmes in Thailand and the Philippines, this literature review drew on findings from previous research on PLH for Young Children (Wessels, 2017) and other parenting programmes in LMICs (e.g. Shenderovich et al., 2018), as well as the substantial body of research available from HICs. Consistent with the focus of the present study, the review only covers factors specifically linked to attendance and the related concept of attrition (i.e. whether participants dropout of the intervention before the last session) rather than examining participation more broadly. Factors understood to effect attendance in parenting programmes were synthesised (albeit not via systematic review) and broadly grouped into four categories: economic and educational; social and health; parenting and child behaviour; and sociodemographic characteristics.

Economic and Educational Characteristics

Among the most frequently examined factors associated with attendance in parenting programmes are socio-economic status and related indices such as employment status and education level. While some studies have demonstrated that lower socio-economic status is

related to lower attendance in parenting programmes (e.g. Peters et al., 2005), several studies found no relation between these two variables (e.g. Axford et al., 2012; Baker et al., 2011; Nix et al., 2009; Reyno & McGrath, 2006). In a review of 262 behavioural parenting programmes in HICs, Chacko and colleagues (2016) found that lower socio-economic status had a moderate negative effect with increased dropout in comparison to those with higher socio-economic status (Cohen's $d = .58$, $p < .010$). Specifically, studies which included participants from lower socio-economic backgrounds had higher dropout rates (34%) compared to studies that included participants from higher socio-economic backgrounds (24%). Similar results were identified in a meta-analysis examining the relationship between household income and attrition in parenting programmes in HICs, which showed that poorer households were significantly more likely to drop out ($r = .21$, $p < .002$) (Reyno & McGrath, 2006).

Regarding attendance at programme sessions, Chacko and colleagues (2016) found no significant association between attendance and socio-economic status in their review (Cohen's $d = -.054$, $p = .100$). However, some individual studies have found some association between these variables. For example, a trial of Parenting Our Children to Excellence for mothers of pre-schoolers ($N = 451$) – a parenting intervention aimed at reducing adverse child behaviour outcomes and improving effective parenting – found that mothers with more time demands and scheduling conflicts attended fewer sessions if they were from average ($t = -5.24$, $p < .001$) and high income families ($t = -4.36$, $p < .001$), but no such association was found for low income families (Dumas et al., 2007). Overall, these different findings may be due to several factors, including a mixture of high-risk and universal samples, different study contexts and interventions, as well as discrepancies in the measurement of socio-economic status.

Some studies have also examined other indicators of socio-economic status including caregiver education, occupational prestige, employment status, household size, and household hunger (e.g. Kazdin, 2000; Nix et al., 2009; Shenderovich et al., 2018; Wessels, 2017). For example, in a study of parent management training delivered within Fast Track (*N* = 445) in the United States, Nix and colleagues (2009) found that parents with lower occupational prestige and lower education levels were equally likely to attend programme sessions as those with higher levels of occupational prestige and education. However, in a meta-analysis of six studies, Reyno and McGrath (2006) found that higher education and occupational prestige was significantly associated with increased dropout ($r = .26$, $p < .001$). A recent study of a fatherhood education programme with 1,040 participants in the United States also found that lower levels of education was significantly related with higher dropout, and that fathers with high school education or less missed approximately 1.9 times more sessions compared to fathers with university degrees (Laxman et al., 2019).

Additionally, the recent PLH for Adolescents trial in South Africa (*N* = 552 caregiver-adolescent dyads) for caregivers and their adolescents aged 10-17 years (Shenderovich et al., 2018) found that caregivers living in peri-urban settings attended approximately 3.08 fewer programme sessions compared to caregivers living in rural villages (estimate = -3.08, $p < .050$). The authors attribute these finding to the fact that rural areas in South Africa are more prone to lower levels of income and access to social services than urban areas. This, in turn, may results in a greater demand for service and thus boosts attendance. Furthermore, while household size was not significantly associated with caregiver attendance, this study also found that employed caregivers attended 3.08 fewer sessions compared to unemployed caregivers (estimate = -3.08, $p < .001$). On the other hand, the PLH for Young Children trial found no significant associations between attendance and household employment (odds ratio = 0.97, 95% CI [0.61, 2.57]) or related indices such as household hunger (odds ratio = 0.92,

95% CI [0.81, 1.05]) (Wessels, 2017). These findings may be attributed to limited variation in socio-economic levels across the sample, as all families were from similar very low-income communities in Cape Town, and may be an artefact of the fact that employed caregivers tended to drop out of the study prior to randomisation.

Social and Health Characteristics

Social and health factors such as caregiver mental health have also been shown to be associated with attendance in parenting interventions. Some studies have demonstrated significant associations between mental health problems, such as higher levels of parental depression and parenting stress, with increased dropout (Calam et al., 2002; Peters et al., 2005). However, most parenting interventions that examined mental health variables found that having more mental health problems was associated with equal or higher attendance (e.g. Baydar et al., 2003; Smith et al., 2018). For example, a meta-analysis showed that neither maternal depression nor parenting stress was significantly related to dropout (Reyno & McGrath, 2006). However, when authors combined parenting stress and negative life events, they found small significant effects associated with increased dropout (r = .15, p = .038).

Additional measures of caregiver well-being that have been examined in relation to attendance in parenting programmes in LMICs include social and health-related variables such as caregivers' experiences of intimate partner violence, caregiver HIV status, and alcohol and substance use. Only two studies that have investigated the association between attendance and intimate partner violence could be identified (i.e. Shenderovich et al., 2018; Wessels, 2017). Neither Wessels (2017) nor Shenderovich and colleagues (2018) found significant effects (odds ratio = .93, 95% CI [0.44, 1.97], and: estimate = -.07, 95% CI [-.47, .34], respectively), suggesting that caregivers experiencing such difficulties at baseline did not attend less than those who were not experiencing difficulties. Shenderovich and colleagues (2018) also examined the association between attendance and HIV status, finding

that there was no significant association (estimate = 0.31, 95% CI [-.76, 1.37]). However, caregivers with higher levels of alcohol and substance use were found to attended 0.60 fewer sessions (p = 0.048) (Shenderovich et al., 2018).

Caregiver and Child Characteristics

There is also evidence, although mixed, that pre-intervention levels of parenting and child behaviour are important factors affecting attendance in parenting programmes. While caregivers with greater parenting difficulties may be expected to find it harder to participate, some studies found the opposite was true. For example, a study of 175 low-income families enrolled in the SAFE Children parenting programme in the United States (Gorman-Smith et al., 2002) found that poor parental monitoring was significantly related to increased attendance in the programme (p = .05). However, other studies have been unable to find this association between parenting difficulties and attendance (e.g. Salari & Filus, 2017; Shenderovich et al., 2018). Additionally, Shenderovich and colleagues (2018) found that caregivers with more positive and involved parenting at baseline attended 0.67 more sessions than caregivers with lower levels of positive and involved parenting (p < .05).

Research on the relationship between pre-intervention levels child behaviour and caregiver attendance has also generated conflicting results. Although numerous studies in both HIC and LMICs (Dumas et al., 2007; Peters et al., 2005; Shenderovich et al., 2018; Wessels, 2017) found no association between baseline child behaviour difficulties and attendance, other studies found that caregivers of children with higher levels of conduct problems attend more programme sessions (Heinrichs et al., 2005; Sirles, 1990). For instance, in a German study of a Triple P parenting programme targeting caregivers with children ages 3-6 years (N = 280), Heinrichs and colleagues (2005) found that caregivers who reported higher levels of externalising child behaviour at baseline were 1.06 times more likely to attend programmes sessions (95% CI = 1.03, 1.12).

Sociodemographic Characteristics

Sociodemographic characteristics, particularly caregiver age and gender, child age and gender, and relationship status, have also been investigated in terms of their relationship to attendance in parenting programmes. However, findings regarding these variables are mostly mixed. In terms of caregiver age, some studies found that being older was related to higher attendance at programme sessions (Nix et al., 2009; Peters et al., 2005; Shenderovich et al., 2018), while others found no association between parental age and attendance (Dumas et al., 2007; Wessels, 2017). A meta-analysis of six studies indicated that maternal age was significantly associated with attrition (r = .21, p = .001), with younger mothers more likely to drop out (Reyno & McGrath, 2006). Similarly, in their study of a fatherhood education programme, Laxman and colleagues (2019) found that while younger fathers were not at an increased risk of dropout, they were significantly less likely to attend programme sessions, with fathers under the age of 25 years missing approximately 1.97 times as many sessions as father over the age of 25 years (p = .013). Although the reason younger parents appear to be less likely to attend is not clear, authors highlighted that younger parents typically have younger children, need to juggle studies and work, and/or are early in their careers leading to more time constraints and schedule conflicts.

Regarding caregiver gender, the samples in most studies of parenting programmes predominantly consist of females; therefore, there is limited research on the association between caregiver gender and attendance. Shenderovich and colleagues (2018) found that male caregivers attended 3.37 fewer sessions than female caregivers, although only eight male caregivers participated in the study. Similarly, a pre-post study of PLH for Young Children in South Sudan (N = 97) which comprised of 23.7% male caregivers found that attendance among female caregivers was significantly higher (standardised $\beta = 0.28$, $p < .001$) than male for caregivers (Janowski et al., 2020). However, this same study found no

significant association between caregiver gender and attendance in a pre-post study of the PLH for Adolescent programme (N = 108) in South Sudan, where 24.3% of the sample were male (standardised β = 0.01, p = .836).

Several studies have also examined the association between attendance and child demographics such as age and gender. While most studies found no association between attendance and child age (e.g. Chacko et al., 2016; Dumas et al., 2007; Wessels, 2017), the study of PLH for Young Children in South Sudan showed that being a caregiver of an older child was significantly related to higher attendance (standardised β = 0.15, p = .025) (Janowski et al., 2020). Regarding child gender, most studies have demonstrated that it was not significantly related to attendance (Dumas et al., 2007; Heinrichs et al., 2005; Wessels, 2017). However, one study of a parenting intervention focussing on adolescent drug abuse treatment (N = 224) found that caregivers with adolescent girls were twice as likely to attended compared to those with adolescent boys (p = .059), although these results only approached significance (Dakof et al., 2001).

Caregiver relationship status may also be associated with attendance in parenting programmes, with single caregivers at higher risk of attrition (e.g. Baker et al., 2011; Reyno & McGrath, 2006). For instance, in their meta-analysis of nine studies, Reyno and McGrath (2006) identified that being a single parent was significantly related to higher levels of dropout (r = .18, p = .005). Researchers have speculated that this may be because single caregivers may have more childcare responsibilities and less time available to attend sessions (e.g. Baker et al., 2011). On the other hand, in their study of 451 low-income mothers enrolled in the Parenting Our Children to Excellence parenting programme in the United States, Dumas and colleagues (2007) found that single mothers were more likely to attend (35%) compared to mothers in a relationship (27%). However, other studies have found that

relationship status was not significantly related to dropout (e.g. Laxman et al., 2019) or attendance (e.g. Laxman et al., 2019; Wessels, 2017).

To conclude, the wide range of factors related to attendance indicate that participation in parenting programmes is not straightforward. No single factor appears to be implicated and the barriers and facilitators that one family faces may be vastly different for another. While this review drew heavily on literature from lower-income settings within HICs, the increased prevalence of social and environmental factors in high-stress, low-resource settings of LMIC, which includes community violence, poor transport systems, and serious health concerns such as HIV, may further influence the ways in which caregivers are able to participate in parenting programmes (Wessels et al., 2016). Since effects of these risk factors in LMICs have almost exclusively been studied in the context of South Africa (e.g. Shenderovich et al., 2018; Wessels, 2017), findings may not be generalisable in other LMICs such as Thailand and the Philippines. Therefore, it is important to investigate the ways in which attendance is affected in such contexts as it will further the understanding of the barriers caregivers face and how those barriers might be addressed.

The Importance of Implementation

At this stage of the implementation, where there are national interests in embedding PLH for Young Children in existing service delivery systems in Thailand and the Philippines, it is vital that implementation processes and their impact on outcomes are investigated. Indeed, evidence from a range of prevention interventions indicates that the quality of implementation achieved is a key factor associated with improvements in programme outcomes. In a review of prevention interventions for young children, 45 (72%) of the 59 studies showed significant positive associations between implementation quality and half or more of the targeted outcomes (Durlak & DuPre, 2008). Similarly, a systematic review of parenting programmes aimed at reducing child behaviour problems among children aged 3-

12 years (Furlong et al., 2013) showed that studies reporting higher levels of implementation produced moderate, statistically significant intervention effects (standardised mean difference = -0.58, $p < .001$), whereas those with lower implementation rates produced small but non-significant intervention effects (standardised mean difference = -0.28, $p = .51$) (Furlong et al., 2013). In addition to providing a theoretical understanding of programme implementation, identifying dimensions that are associated with positive outcomes have important implications for developing and informing monitoring and evaluation systems to maintain programme effectiveness within routine practise (Berkel et al., 2011; Berkel et al., 2018). These have been identified as essential strategies for improving translational research which attempts to bridge the gap between the carefully controlled conditions of trials and effective implementation within real-world settings (Berkel et al., 2018).

Programme implementation has been conceptualised as a construct consisting of four distinct components, namely: quality, adaptation, fidelity, and participant responsiveness (Berkel et al., 2011). Participant responsiveness, defined as the level of interest and involvement in the programme by participants (Dane & Schneider, 1998), has been operationalised using four indices: session attendance, quality of participation during sessions, home practice of skills learnt during sessions, and programme satisfaction (Berkel et al., 2011). Several studies investigating the influence of implementation on outcomes have found that composite measures of responsiveness are significantly associated with programme outcomes (e.g. Schoenfelder et al., 2013). Indicators of responsiveness have also been considered individually.

A number of studies have investigated the association between attendance and programme outcomes. Gross and colleagues (2009), in their study of the Chicago Parent Program ($N = 253$) targeting low-income African American and Latino caregivers of children aged 2-4 years who were enrolled in day care centres in Chicago, found that caregivers who

attended at least 50% of the programme sessions reported higher increases in parenting self-efficacy (Cohen's $d = .37$, $p < .05$), more consistent discipline (Cohen's $d = .29$, $p < .01$), less corporal punishment (Cohen's $d = -.30$, $p < .01$), more warmth (Cohen's $d = .08$, $p < .05$), and greater reductions in child behaviour problem intensity (Cohen's $d = -.31$, $p < .05$) when compared to the control group. These results align with several meta-analyses (Menting et al., 2013; Reyno & McGrath, 2006; Wilson & Lipsey, 2001). For example, a meta-analysis of 51 studies of Incredible Years programmes by Menting and colleagues (2013) found that higher attendance at sessions was positively related to programme effects when controlling for severity of child problem behaviour at baseline ($\beta = .37$, $p = .04$). Similarly, Reyno and McGrath (2006) found that attendance at sessions was significantly associated with intervention outcomes, although the effect was small ($r = .16$, $p = .04$).

However, not all studies have found attendance to be associated with positive outcomes. While a recent study of an Incredible Years programme of 387 caregiver-child dyads in the Netherlands showed that higher attendance at sessions was associated with less caregiver-reported negative parenting behaviour ($\beta = -.012$, $p < .001$), and more observed positive parenting behaviour ($\beta = .018$, $p = .001$), higher attendance did not impact effects on reported or observed child externalising behaviour, prosocial child behaviour, observed negative parenting behaviour, and reported positive parenting behaviour (Weeland et al., 2017). In addition, in their study of a parent management programme delivered within Fast Track, Nix and colleagues (2009) found that attendance at sessions was not associated with change in any parenting outcomes, including parental warmth, parental perceptions of their child, school involvement, and physical punishment. However, this same study found that active participation in sessions predicted improvements in parental warmth ($\beta = .14$, $p < .100$), parental perceptions ($\beta = .14$, $p < .001$), school involvement ($\beta = .23$, $p = .001$), and physical punishment ($\beta = -.16$, $p < .001$), suggesting that attending without actively engaging

may not be enough to impact outcomes. In contrast, the recent PLH for Adolescents programme in South Africa showed that neither attendance or active participation were significantly associated with programme outcomes (Shenderovich et al., 2019).

Accounting for Attendance When Estimating Programme Effects

Although evidence regarding session attendance and its impact on outcomes is inconclusive, there appears to be a dose-response relationship for some, if not all outcomes in most of the reviewed studies. This suggests that treatment allocation (randomisation) influences the outcome via the amount of treatment (i.e., attendance) received. Thus, treatment allocation should, theoretically, be strongly correlated to attendance and only influence outcomes through attendance (Dunn et al., 2005; Stuart et al., 2008). However, the effect of attendance on outcomes is likely confounded by measured and unmeasured variables (Angrist et al., 1996). For instance, variables such as baseline child problem behaviour or structural barriers such as socio-economic status might be related both to the number of sessions attended and to the programme outcomes. As a result, the dose-response effect, or the causal effect of attendance, will be confounded. While this is not an issue for intent-to-treat (ITT) approaches which do not account for attendance, establishing causal effects in dose-response approaches can be problematic due to the possible confounding effects of moderators.

ITT analyses typically compare intervention group outcomes to those of the control group regardless of whether they dropped out or whether they received the allotted programme dosage (i.e., attendance) (Huang et al., 2014). In doing so, this approach retains the integrity of randomisation and provides relatively unbiased effects for intervention assignment and study dropout (Huang et al., 2014; Stuart et al., 2008). An advantage of this approach is that by aggregating intervention effects for varying levels of attendance, it may model real-world treatment effects more closely (Huang et al., 2014). That is, if attendance

variations exist within a trial context, they will likely also exist if the programme is implemented more broadly. However, a limitation of ITT analyses is that it does not account for low dosage or poor attendance – also referred to as noncompliance (Jo & Muthén, 2001). For example, although parenting programmes usually comprise of multiple programme sessions, some caregivers assigned to the intervention group may only attend a few of these sessions or may not attend any sessions at all. Since ITT analyses do not account for varying levels of attendance among participants, they tend to underestimate the magnitude of programme effects (Jo & Muthén, 2001). Thus, there have been increasing calls to account for implementation factors such as participant attendance when estimating programme effects from randomised controlled trials (Stuart et al., 2008).

To supplement ITT estimates, as-treated and per-protocol approaches are frequently used to estimate the association between attendance and outcomes. However, as-treated analyses typically group participants by the amount of intervention received rather than by treatment assignment (i.e., participants who received less than the required dosage are grouped with the control group), while per-protocol analyses group participants by treatment assigned but remove noncomplying participants from the analyses (i.e., participants who receive less than the required dosage are dropped from the analysis but only for the intervention group) (McNamee, 2009). Besides not accounting for confounding moderators, a significant shortcoming of these approaches is that they introduce even more bias by "breaking" randomisation and removing noncompliant participants from the sample (Little & Yau, 1998; Sedwick, 2015). One analytical approach that overcomes these challenges is Complier Average Causal Effect (CACE) analysis and the related instrumental variable method. This approach extends Rubin's causal model and provides estimates of programme effects using data from compliers and noncompliers across both arms of the intervention,

thereby producing less biased programme effect estimates than as-treated or per-protocol analyses (Angrist et al., 1996).

The estimation of CACE is commonly implemented using a two-stage least squares algorithm (Stuart et al., 2008). Two-stage least squares involves two regression models, first regressing attendance on the "instrument" (randomisation to intervention or control group), and then regressing the outcome on the predicted value of attendance which was obtained from the first regression (Dunn et al., 2005). The "instrument" (treatment allocation) is assumed to influence attendance but not programme outcomes (except through its association with attendance). Therefore, the treatment indicator variable is only used in the first model of attendance, but not in the second which models outcome thus accounting for measured or unmeasured confounding in the sample.

In contrast to ITT analyses which estimate programme effects for the entire study regardless of dropout or dosage, CACE estimates are useful for estimating the "pure" programme effects for participants who received above or below a certain dose of the programme (Schochet & Chiang, 2011). This can be especially beneficial for efficacy studies where the focus is on determining whether the intervention worked. Programme implementers may also find CACE estimates useful to inform and improve implementation efforts. For instance, knowing what programme effects could be expected from specific attendance rates might help implementing organisations devise strategies to achieve these rates. Furthermore, CACE estimates can also be used to draw important policy lessons from ITT effects (Schochet & Chiang, 2009, 2011). For example, CACE effects can help determine whether weak or null ITT effects are due to small programme effects among participants or due to low rates of attendance/compliance.

Despite being a common analytical tool in fields such as medicine, only a handful of studies have used CACE in family-based prevention interventions (e.g. Huang et al., 2014;

Stuart et al., 2008) and even fewer in parenting programme research. Despite an extensive (yet not systematic) search, I have been able to identify only four published studies using CACE in the context of parenting programmes (DeGarmo & Jones, 2019; Kim et al., 2018; Kogan et al., 2016; Stanger et al., 2011). However, these studies were all conducted in HICs where potential moderators of attendance and outcomes are likely to differ substantially from those in LMICs due to the distinctly different social, economic, service delivery contexts in which the programmes are implemented. Conducting CACE analysis can therefore be an important tool for understanding the mechanisms through which engagement in PLH for Young Children affects parenting and child outcomes in the two delivery contexts of Thailand and the Philippines. This is especially important for guiding implementation adaptations in follow-up research and for informing systems to monitor and maintain effectiveness once these programmes are implemented more routinely.

In summary, studies suggest a wide range of factors related to attendance in parenting interventions. However, findings have mostly been inconsistent and studied in the context of HICs. Furthermore, while a number of parenting trials have investigated the influence of attendance on programme outcomes, these studies predominantly used relatively biased dose-response approaches which tend to break randomisation, remove noncompliers from the analyses, or fail to account for confounding effects. CACE analyses, which avoid these challenges, have been used to estimate programme effects in parenting interventions, but also only in the context of HICs. This study therefore sought to address these two crucial gaps in the literature.

Research Aims and Hypotheses

Given the mixed findings regarding predictors of attendance and the paucity of research form LMICs, the first aim of this study was to explore what baseline characteristics predict attendance in the two randomised controlled trials of PLH for Young Children, one in

Thailand and another in the Philippines. Although the variables used as potential predictors were based on the literature review and the available data in each country, this study did not hypothesise the direction of associations with attendance due to the inconsistency of previous findings. Baseline factors used as prospective predictors of attendance included economic and educational characteristics (family income; government benefits; household assets; household size; household hunger; caregiver education), social and health characteristics (caregiver physical health; caregiver mental health; caregiver history of abuse; intimate partner violence), parenting and child behaviour (child maltreatment; positive parenting; dysfunctional parenting; poor monitoring and supervision; child behaviour problems; parent daily report on child behaviour and parenting), and sociodemographic characteristics (caregiver age; child age and gender; relationship status; caregiver employment).

The second aim of this study was to examine how programme effects of PLH for Young Children in Thailand and the Philippines may be impacted by attendance variability. This aim will be achieved by comparing traditional ITT estimates to CACE estimates to evaluate the effects of the programmes at post-test and follow-up on the primary outcomes of child maltreatment (physical and emotional abuse and neglect) and on the secondary outcomes of positive parenting; dysfunctional parenting; child monitoring and supervision; child behaviour problems; and caregiver mental health.

The specific hypotheses that were tested are as follows:

1. Baseline characteristics (e.g. economic and educational; social and health; parenting and child behaviour; and sociodemographic) will be associated with caregiver attendance among participants allocated to the PLH for Young Children programme in Thailand and the Philippines.
2. Higher session attendance by caregivers enrolled in the PLH for Young Children programme in Thailand and the Philippines will be associated with greater

improvements in the primary outcomes of child maltreatment and the secondary outcomes of positive parenting; dysfunctional parenting; child monitoring and supervision; parental mental health; and child behaviour problems.

Chapter 2: Method

This study is a secondary data analysis of two RCTs of PLH for Young Children, one implemented in Thailand and the other in the Philippines. Before discussing the methods used to address the aims of the present study, this chapter first provides contextual background to the various components of the RCTs, including study settings, procedures, data collection, adaptation and feasibility piloting of the programmes, programme delivery, and ethical considerations. The second half of this chapter covers the specific measures used for the secondary data analyses in this thesis, as well as the analytic strategies employed to answer the two study hypotheses.

Study Setting

Thailand

Data collection and programme delivery was conducted in Udon Thani, one of 20 provinces in the North Eastern region of Thailand. This study site was chosen, in collaboration with the Thai Ministry of Public Health and UNICEF Thailand, and the Udon Thani Muang District Public Health Office, due to the lack of available social services in this region, as well as the well-established working relationship between national and local government officials and UNICEF Thailand. Udon Thani has a total population of 996,128 people (Provincial Community Development Office of Udon Thani, 2017) and is characterised by high rates of income inequality (Gini coefficient of 0.53) and poverty (United Nations Development Programme, 2010). Although 93.4% of households own land in the province, households earn an average monthly income of 17,273 Thai Baht (≈ USD 548,87) with income presumed to be even lower in rural areas (National Statistical Office Thailand, 2015). The province also has low rates of formal education, with an average of 7.6 years of schooling (United Nations Development Programme, 2010). The parenting programme was delivered in Health Promotion Hospitals (i.e., local community clinics) by

public health officers and nurses and recruited participants were primarily from low-income, peri-urban, and rural communities.

Philippines

Programme implementation took place in Western Bicutan, an urban and densely populated district in Taguig City. Taguig is one of 16 highly urbanised cities in the National Capital region of the Philippines. Western Bicutan has a population 91,158 people, making up 11.3% of the total Taguig population (Philippine Statistics Authority, 2015). While there are no available data on income and poverty rates for Western Bicutan itself, the 2018 Family Income and Expenditure Survey indicates that the average household income in the National Capital region is 38,333 Philippine Pesos per month (≈ USD 779,72), and the unemployment rate is approximately 12.3% (Philippine Statistics Authority, 2018). The primary reason for selecting Western Bicutan as the study site was that the Pantawid Pamilyang Pilipino Programme (4Ps), a conditional cash transfer programme for low-income families implemented by the Philippine Department of Social Welfare and Development (DSWD), is delivered to over 2,000 families in this community. The DSWD wished to test the parenting programme for wider roll-out within the 4Ps system. Western Bicutan was selected as an appropriate site for the RCT due to the high numbers of 4Ps beneficiaries and it being a priority area for the DSWD and UNICEF Philippines, one of the main funders of the study. The target age of 4Ps families with children aged 2-6 years was due to constraints imposed by another donor, the UBS Optimus Foundation, which focused on early years of child development only.

Study Procedure

Participant Recruitment

Participants were recruited using targeted sampling – a form of nonprobability sampling - to identify caregivers of young children who experienced risks of maltreating their

children. Both RCTs screened for low-income families, but families with younger children were also targeted given the high rates of violence with this age group. The recruitment process was implemented through local partners who identified potential eligible caregivers and referred them to the research team. In Thailand, primary caregivers were identified through local Health Promotion Hospitals, community health volunteers working in the targeted areas, and referrals from local teachers in the region. In the Philippines, the research team worked closely with local government staff to identify potential caregivers who were recipients of the conditional cash transfer system. Once identified, caregivers were invited to participate in a community meeting in which the study was explained to them and to determine whether they were interested in participating in the programme. Rolling recruitment and enrolment strategies were used in both countries, with recruitment ending once 120 caregivers had successfully enrolled in each study.

Overview of Data Collection Procedures

In Thailand, 10 data collectors who were fluent in Thai (the local language in Thailand) and who had prior experience working with low-income families were recruited by the Boromarajonani College of Nursing in Udon Thani through the Regional Health Centre. Similarly, 10 data collectors fluent in Filipino were recruited by the Bulatao Centre at the Ateneo de Manila University in the Philippines. Prior to data collection, recruited data collectors participated in a 20-hour training workshop conducted by Dr Lachman (Oxford) and Professors Liane Peña Alampay (Philippines) and Frances Gardner (Oxford). These workshops provided training in research ethics, informed consent procedures, child safety and protection procedures, interviewing skills, observational assessment techniques, and how to manage disclosures of potential or actual harm against children or other participants.

Screening questionnaires, informed consent, outcome questionnaires, and observational assessment (in Thailand only) were administered using Computer-Assisted

Self-Interviewing (CASI) methods with electronic tablets. This approach to data collection was piloted during the feasibility studies in both Thailand and the Philippines and was found to be highly acceptable by participants.

CASI also allows for an audio-enhanced function, known as audio-CASI, in which respondents can play audio recordings of the questions. Audio-CASI was used for sensitive items such as those concerning child maltreatment and intimate partner violence. In a menta-analysis of quantitative interviewing tools in LMICs, audio-CASI was found to increase participants' disclosure of sensitive or stigmatising activities and/or experiences compared to face-to-face interviews (Phillips et al., 2010). Audio-CASI may help lower participant's anxieties in responding to face-to-face questions, thereby increasing disclosure of maltreatment (Connolly, 2005; Davies & Morgan, 2005). Furthermore, this approach may also help address low literacy levels common among low-income and older populations, such as those in both countries. Participants therefore had the option of reading or using the audio function when answering questions.

Eligibility Criteria

CASI electronic tablets were used to administer the screening survey to determine study eligibility of all recruited participants. In cases where caregivers had more than one child in the targeted age range (see eligibility criteria below), they were asked to select the child who had the most difficult behaviour to manage (Thailand) or CASI tablets were used to randomly select one of the children for the caregiver to report on throughout the study (Philippines). To be eligible for study inclusion, caregivers and their targeted child had to meet the criteria listed below.

Caregiver inclusion criteria:

1. Male or female aged 18 years or older;

2. Serves as the primary caregiver of a child between the ages of 2-9 years (Thailand) or 2-6 years (Philippines);
3. Lives in the same household as the target child for at least four nights per week;
4. Agrees to participate in the PLH for Young Children parenting programme;
5. Is a recipient of the Pantawid Pamilyang Pilipino Programme (4Ps) conditional cash transfer programme (Philippines only);
6. Provides consent to participate in the study; and
7. Completes the baseline assessment.

Caregiver exclusion criteria:

1. Any caregiver with acute mental disabilities or severe mental health problems;
2. Any caregiver who was not available to participate in the programme at the time of recruitment; and
3. Any caregiver who had already participated in the feasibility pilot study (a feasibility pilot study was conducted in each country to assess suitability of the programme for these contexts – more details are provided below).

Child inclusion criteria:

1. Aged 2-9 year in Thailand or 2-6 years in the Philippines; and
2. Primary caregiver meets the inclusion criteria listed above.

Child exclusion criteria:

1. Any child with severe mental or physical disabilities.

Of the 126 participants that were screened for eligibility in Thailand, 120 were eligible to participate and were contacted for informed consent. In the Philippines, 124 participants were screened, and 120 were also found to be eligible and consented to participate.

Baseline Assessment

Outcome questionnaires were administered to participants at three time points, the first of which was prior to the start of the programme (i.e., at baseline). For each RCT, baseline questionnaires were administered to participants in their homes and lasted approximately 60-90 minutes. Before baseline assessment, trained data collectors explained how to use the electronic tablet, as well as read out questionnaire questions and assisted participants with filling in their responses. Participants also received a tutorial on how to use the audio function to answer sensitive items and were provided with earphones to ensure privacy. Pen-and-paper interviewer-assisted questionnaires were also available for any participants were uncomfortable or unable to complete the electronic version.

In addition, in Thailand data collectors also conducted observational assessments using an adapted version of the HOME Observation Assessment (Bradley & Corwyn, 2005), in the homes of all participants, which lasted between 30 and 60 minutes. To ensure accurate assessment, data collectors received additional training including in practicing assessments according to a scripted guide. Observational assessments were recorded on the tablets using a checklist format.

Randomisation

The caregivers who participated in the baseline assessment (N = 120 in Thailand, N = 120 in the Philippines) were randomised on a 1:1 allocation ratio. In the Philippines, control group participants received an active comparison treatment as part of the standard monthly Family Development Services that are part of the 4Ps programme. In Thailand, participants were allocated to a services-as-usual non-active condition. In both countries, the randomisation sequence was generated by external researchers based at the University of Oxford using the concealed computerised programme SealedEnvelope™. The research team only notified participants of their allocation status once the baseline assessment had been

completed to ensure that caregivers remained blind to their allocation during the baseline assessment. Although blinding participants to their own allocation status was not possible thereafter, the allocation status of other participants (participants not in their parenting group) remained concealed to participants to lower risk of contamination. Data collectors and statisticians were also blind to allocation.

Intervention

The original PLH for Young Children intervention is a 12-session parenting programme delivered by trained facilitators to groups of caregivers with children aged 2-9 years. The programme uses a metaphor of building a 'House of Support' where the walls of the house represent positive parenting and the roof represents limit setting and non-violent discipline strategies. Facilitators use a collaborative nondidactic approach to build parenting skills and sessions cover spending quality time with children, naming actions and feelings, using praise and rewards, giving instructions, establishing household rules, using non-violent discipline techniques, and problem-solving. The final session reviews what caregivers have learned and ends in a celebration. This programme was adapted, and pilot tested for delivery within Thailand and the Philippines following the guidelines on developing and evaluating complex social interventions set out by the UK Medical Research Council (Craig et al., 2008).

First, formative evaluations in Thailand and the Philippines examined the cultural and contextual relevance of the PLH for Young Children programme themes, structure, process, and schedule and logistics for delivery in a Thai and Filipino context. The formative evaluation in Thailand involved consultation with government officials, practitioners, and professionals from Udon Thani, Bangkok, and Chiang Mai city through focus group discussions and in-depth interviews. The findings and recommendations from these discussions and interviews were then used to inform a subsequent feasibility pilot (N = 30

parents) of an eight-session version of the programme. Programme content and delivery were finalised after consolidating results from the feasibility pilot study. Similarly, a formative evaluation in the Philippines used focus group discussions and individual interviews with service providers and caregivers to examine the cultural relevance and appropriateness of the programme content and mode of delivery of the PLH for Young Children in Western Bicutan and the Philippines more broadly. Based on the findings of this evaluation, PLH for Young Children was adapted and piloted in a small feasibility study (N = 30) to further assess the feasibility of the adapted programme. This version was specifically adapted for caregivers and their children aged 2-6 years due to donor constraints.

Both programmes had similar content: 1) spending quality time with children; 2) describing children's feelings and actions for socio-emotional awareness and cognitive development; 3) using praise and rewards to encourage positive behaviours; 4) setting limits through consistent daily routines and household rules as well as through instruction giving; 5) ignoring negative attention seeking and demanding behaviours, 6) non-violent consequences for rule-breaking and noncompliance; and 7) mindfulness based stress reduction for caregivers. Core activities during sessions included practising skills, problem solving, illustrated stories, discussions regarding specific parenting principles, and discussion about assigned home activities.

In addition, surface level adaptations of caregiver manuals, which participants use as a reference during home practice, were conducted to improve accessibility and acceptability of content. This primarily involved redrawing illustrated stories to make them appropriate for a Thai and Filipino audience. Programme content was also simplified to include more pictures. The adaptation process was supported by Clowns Without Borders South Africa, a non-government organisation that leads the capacity building and dissemination of the PLH programmes.

Facilitator Training and Supervision

The facilitator manuals, developed in the respective formative evaluation studies, were used to inform the core content for training the facilitators. Dr Lachman, one of the developers of PLH for Young Children and a co-principal investigator on both studies, trained the facilitators alongside local PLH trainers in each country.

Eight facilitators who had previous experience in delivering health services to low-income families and children were trained to deliver the programme in each country. The trained facilitators in Thailand included nurse practitioners, social workers, and mental health officers, whereas trained facilitators in the Philippines comprised of psychology master's students, child protection officers, and community volunteers.

Training took place over a five-day workshop which provided: 1) an overview of programme content and specific parenting skills; 2) techniques for managing group sessions; 3) using collaborative and non-didactic facilitation; 4) conducting collaborative problem solving with participants; and 5) facilitating interactive games and activities in programme delivery. Supervisors were also trained (two in Thailand and two in the Philippines) who provided facilitators with ongoing supportive supervision throughout programme delivery. All facilitators underwent a certification process which involved observational assessments of parenting sessions in order to assess quality of delivery using a PLH-Facilitator Assessment Tool (Mackenzie et al., 2019).

Programme Delivery

PLH for Young Children Thailand. Two facilitator pairs delivered the programme to groups of 15 caregivers over eight weekly sessions. Weekly sessions took place at four Health Promotion Hospitals in Udon Thani. Each group session lasted between 2 – 2.5 hours and covered the content outlined above. Programme attendance data was collected via attendance registers at each session.

To enhance attendance and engagement in the programme, facilitators also delivered home visits to caregivers who were unable to attend group sessions or who were having difficulties with programme content and needed additional support. Receiving a home visit was classified as having attended the session. Home visits covered the same content as the group sessions but in an individualised format. In addition, four weekly text messages were sent to all participants to encourage attendance and home practice. Caregivers also received a 10-minute telephone consultation with a facilitator between every session.

PLH for Young Children Philippines. Trained facilitator pairs delivered the 12 programme sessions every second week to groups of 15 caregivers. Sessions took place in community centres in Western Bicutan, with each session lasting 2 – 2.5 hours. Facilitators administered attendance registers at each session to track participant attendance throughout the programme. Home visits were not provided for caregivers who missed a group session. However, to encourage attendance and to remind caregivers to practise activities learnt in the previous session, six text messages were delivered in between each session to each caregiver throughout the programme. All caregivers also received a 10-minute telephone call from a facilitator between each of the 12 sessions.

Post-Intervention and Follow-Up Assessment

Before post-intervention and follow-up assessment, data collectors reminded participants how to use the electronic tablets. As with the baseline assessment, interview questions were read out loud by data collectors to assist participants with filling in their responses, and participants received earphones to answer sensitive items using the audio-CASI function. Pen-and-paper interviewer-assisted questionnaires were also available if any participants were unable to complete the electronic version.

For both RCTs, outcomes questionnaires were administered to participants in their homes but at slightly different timepoints. In Thailand, data collectors returned to

participants' homes one-month post-intervention to collect post-test data and then again three months post-intervention to collect follow-up data. In the Philippines, data collectors collected post-test data immediately after the programme, with follow-up data collection occurring 12-months post-intervention. Completion of the questionnaires took between 60 to 90 minutes, with an additional observational assessment in Thailand lasting between 30 and 60 minutes.

In Thailand at post-test, four participants (one in the intervention and three in the control group) were lost to attrition, resulting in a sample of 116 participants. An additional two participants dropped out at follow-up (one in the intervention group and the other in the control group) leaving a sample of 114. In the Philippines, four participants also dropped out at post-test (two in the intervention group and two in the control group) with an additional two (one in the intervention group and one in the control group) dropping out at one-year follow-up. Participant flow diagrams summarising the study procedure of PLH for Young Children Thailand and PLH for Young Children Philippines are presented in Figure 1 and Figure 2, respectively.

Figure 1

PLH for Young Children Thailand CONSORT Flow Diagram

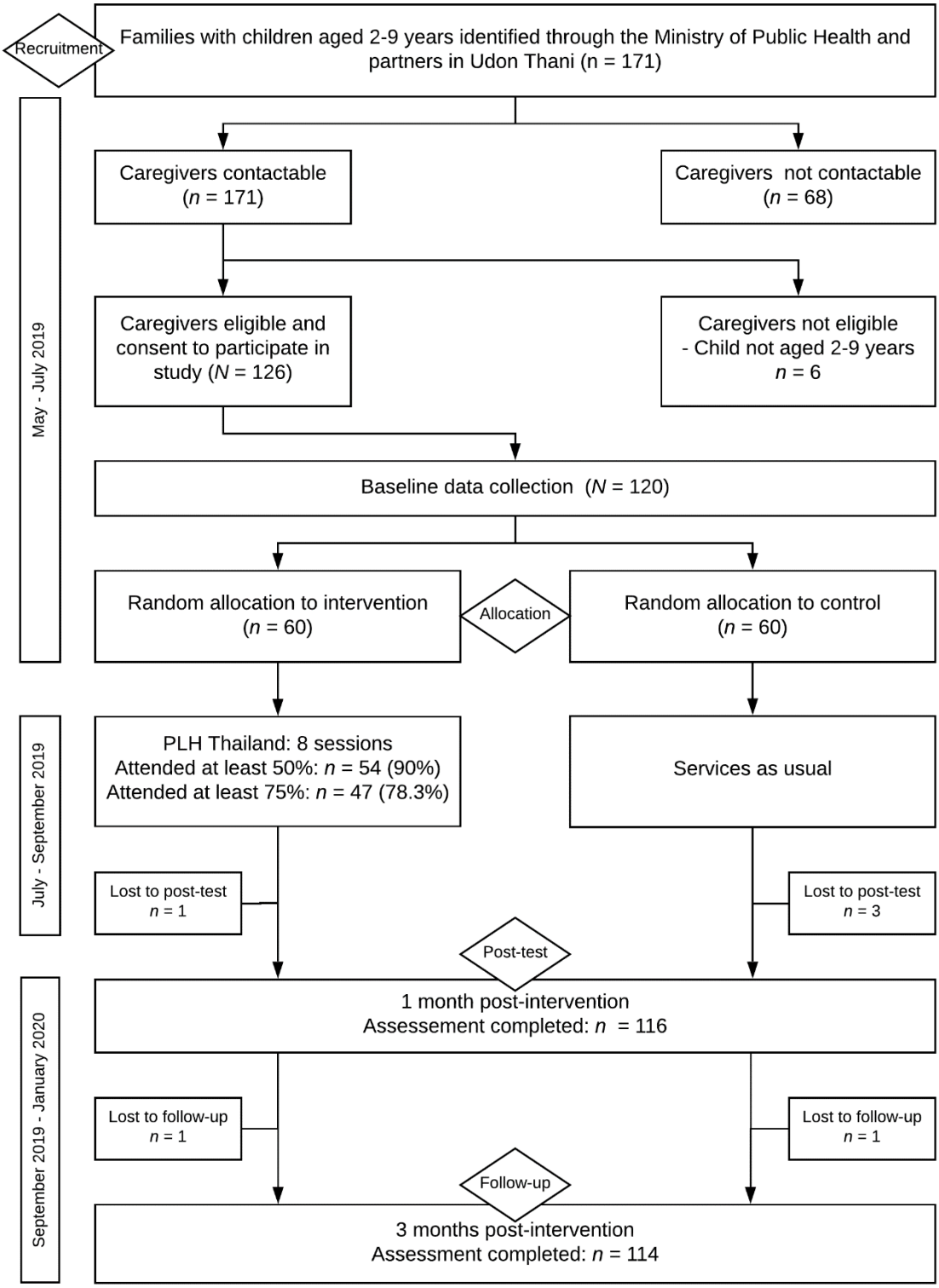
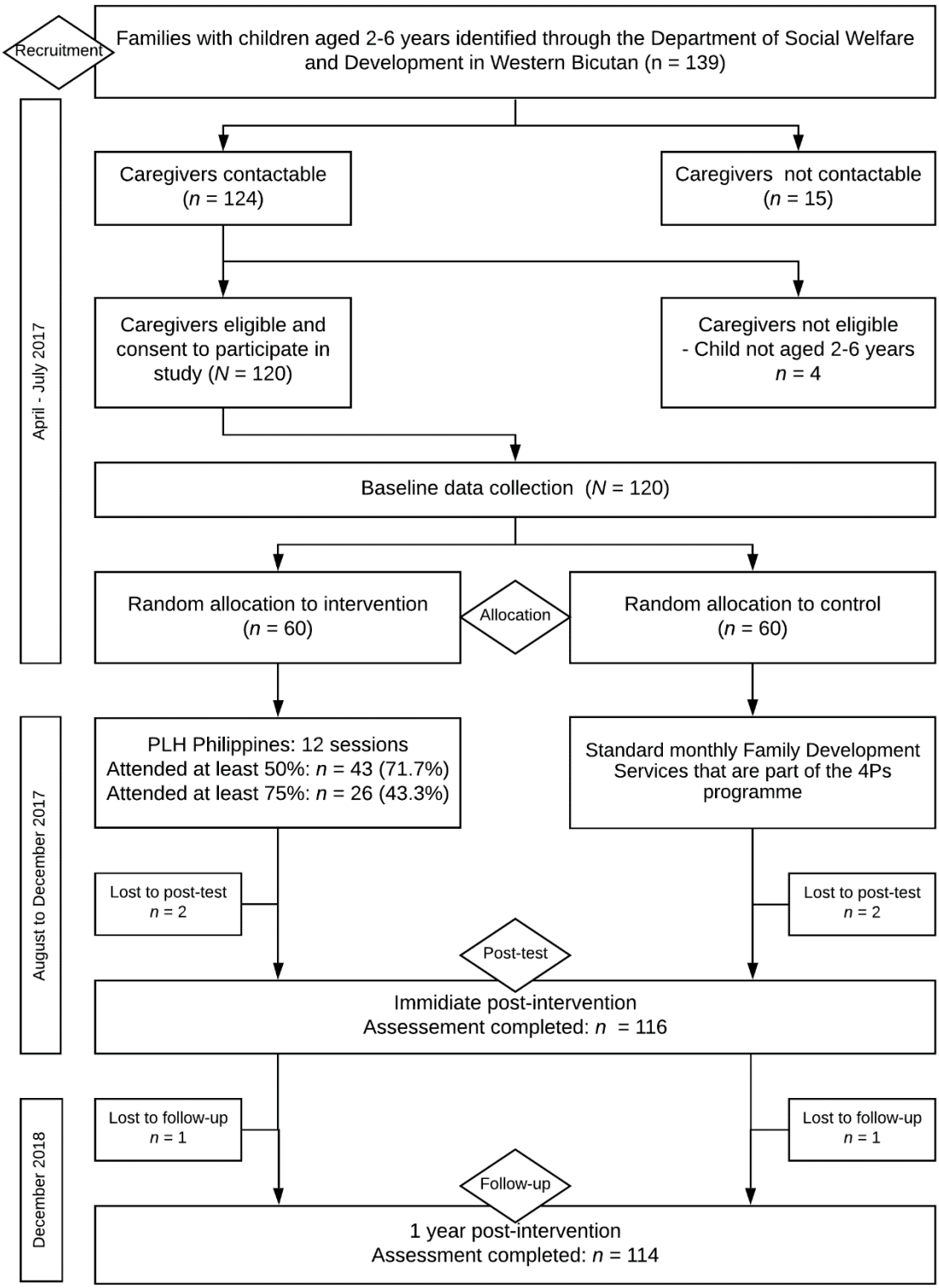


Figure 2

PLH for Young Children Philippines CONSORT Flow Diagram



Ethical Considerations

Ethics Approval

Ethics approval for the RCTs was obtained from institutional review boards and/or research ethics committees within each country as well as from the University of Oxford. For the RCT in Thailand, ethical approval was granted by the ethics review board of the Ministry of Public Health (reference number: 10/2561) and the University of Oxford Tropical Research Ethics Committee (OxTREC ref: 8-18). For the RCT in the Philippines, ethics approval was obtained from the Ateneo de Manila University's Research Ethics Committee (Reference number: AdMUREC-16_90) and the University of Oxford Central University Research Ethics Committee (Reference number: R43041/RE002). Copies of the ethics approval letters are included in Appendix A to D. Both studies were also registered on ClinicalTrials.gov (Thailand: NCT03539341; Philippines: NCT03205449).

Informed Consent

Trained data collectors, who were closely supervised by the respective research teams, conducted informed consent procedures with participants either at their homes or at community centres. Participants received a written copy of the information sheet in the local language (i.e., Thai, or Filipino). Information sheets were also read out loud by data collectors to ensure that participants understood the purpose and process of the study. Data collectors emphasised that participation in the programme and in the study was completely voluntary and that participants could drop out at any time without any negative consequences. Participants were also informed about the specific procedures taken to guarantee their confidentiality during data collection, storage, and analysis (these are presented below). Participants were also informed that everything they said would remain confidential unless they were at risk of harm or of putting someone else at risk of harm.

Participants were given the chance to consent to participate immediately or could wait one week to decide whether they wanted to participate in the study. Consenting participants and the data collector signed the Participant Consent Form on the same date prior to the baseline assessment. Illiterate participants indicated their consent with an "X" or a thumbprint on the consent from.

Confidentiality

Precautions were taken to safeguard the confidentiality of all participants. Each participant was allocated an individual identification number to ensure that their names were not disclosed on any research materials including datasets. Furthermore, all assessment data were anonymised before statistical analysis. Anonymisation included the removal of all personal information that could lead to identification, such as birth dates, telephone numbers, postal addresses, pictures, information of location, and occupation.

All data collected via electronic tablets were encrypted and only made accessible to senior study investigators. Access to the tablets was password protected, and each tablet had a GPS tracking application activated to allow for remote deletion of data in case tablets were stolen. Electronic tablets were also stored in locked cabinets at each study site when they were not in use.

To further ensure confidentiality and data safety, two methods were used to upload and securely store data at the end of each day of data collection. Encrypted data were first transmitted to a secured, password-protected server housed at the University of Oxford. Then, the Research Manager in each country manually uploaded digital data onto a local server which served as a backup to the central server system at the University of Oxford. All data were erased from the tablets after the completion of the study.

Disclosure of Harm or Potential Harm

Caregivers may disclose harsh parenting practises that reflect potential or actual abuse or neglect of children to the data collectors or other members of the research team. To mitigate or respond to actual or potential harm to children during the study, a child protection protocol was followed specifying referral procedures to child welfare, health organisations, and other services. All staff received additional training from the research team on how to respond to these types of situations in alignment with the study's referral protocols. Weekly supervision meetings with all field interviewers were also held to allow for discussion on issues that arose concerning harm to caregivers and children.

Potential Risk of Harm from the Intervention

While this study did not involve processes likely to cause harmful effects, it was possible that participants might have become distressed during data collection due to the sensitive content of the assessment. All research personnel therefore had experience working with vulnerable families and were trained in the ethical protocols and procedures regarding research with human subjects. In cases where participants needed or wanted to access additional support such as attending a clinic or seeing a counsellor, referrals were made by the research team. In addition, participants were provided with self-referral documents describing available services in the area as part of the informed consent procedure as well as at each subsequent data assessment stage. During the consent procedure, participants were informed that everything they said would remain confidential unless they were at risk of harm or of putting a child at risk of harm. Research personnel also signed a confidentiality agreement to protect any disclosure of identifying information about participants. Lastly, participants were informed that they did not have to answer any questions that they felt unsure or uncomfortable with and that they could stop the data assessment or interview without any negative consequences.

Participant Incentives

Offering monetary incentives to participants to further enhance retention is somewhat contentious. Some researchers have expressed concerns that incentives compromise voluntary participation and may also lead to participants withholding information that may deem them ineligible for study participation (e.g. Head, 2009). On the other hand, incentives may also serve as compensation for participants' time and involvement invested in the study and are thus commonly provided in studies of prevention interventions (e.g. Cluver et al., 2018; Ward et al., 2020).

All participants in the Thailand RCT were provided with 150 Thai Baht (≈ USD 4.80) after completing the outcome questionnaire at each data collection point: baseline, one-month post-intervention, and three-months post-intervention. An observational assessment was also administered at each time point, for which participants received 50 Thai Baht (≈ USD 1.60). Additionally, participants received 50 Thai Baht (≈ USD 1.60) for successfully completing each Parent Daily Report which was administered on five occasions (baseline, one month post-baseline, two months post-baseline, one month post-intervention, and three months post-intervention).

In the Philippines, participants received incentives in the form of a grocery voucher (valued at approximately 10 USD) after successfully completing the baseline, post-intervention, and follow-up assessments. They also received a monetary incentive of approximately 5.2 USD for completing all five of the Parent Daily Report assessments. However, it should be noted that participation in the PLH for Young Children programme was not conditional for receiving grants and participants could revert at any time back to services as usual.

Participants in the intervention group did not receive monetary incentives for attending programme sessions. However, caregivers in Thailand were provided with

childcare and were reimbursed for transportation costs to and from group sessions (150 Thai Baht; approximately 4.80 USD). In both RCTs, participants in the intervention group also received meals and light refreshments at group sessions. Participants also received a certificate upon completing the programme. Additionally, all participants who left the study before programme completion also received a certificate stating how many sessions they attended.

Measures Used for Secondary Data Analysis

The following section details the measures used to conduct the analysis for this thesis. Since this is not the exhaustive list of outcomes assessed in the original studies' questionnaires, summaries of the full outcome questionnaires are presented in Appendices E and F. Measures were translated into Thai and Filipino and back-translated into English to confirm the accuracy of the translation. All outcomes and their respective measures are listed in Table 1.

Due to the exploratory nature of this study, baseline variables selected as potential predictors of attendance were chosen based on the literature on factors associated with attendance in parenting programmes in LMICs as well as in HICs. Primary and secondary outcome were used both as baseline predictors in the analyses predicting attendance, as well as outcome variables in the ITT and CACE analyses.

Baseline Household and Family Characteristics

Caregiver and Child Demographics. Basic demographic factors were measured at baseline using several items from the UNICEF Multiple Indicators Cluster Survey (MICS) Household Survey (UNICEF, 2005). The MICS has been used extensively throughout LMICs including Thailand and the Philippines. It was used to assesses caregiver and child age and gender as well as caregiver education, marital status, and relationship with child.

Caregiver Health. Caregiver general health was assessed using one item from an adapted version of the Medical Outcomes Study Short Form-12 Health Survey (Ware et al., 1996). This item assesses caregiver overall physical and mental health on a 5-point Liker-like scale (1 = Poor; 3 = Excellent).

Caregiver Disability. In addition to general health, caregiver disability was also assessed due to the high number grandparents in the Thailand study. This was measured using one item adapted from the Washington Group short set of questions, originally developed for use in surveys and censuses according to the Fundamental Principles of Official Statistics (Madans et al., 2011). This item asks respondents whether they have any difficulties "hearing, seeing, communicating, or moving," and was scored using a categorical variable (0 = No; 1 = Yes).

Caregiver History of Child Abuse. An adapted version of the International Society of Prevention of Child Abuse and Neglect (ISPCAN) Child Abuse Screening Tool Retrospective version (ICAST-R) was used to measured caregivers' own experience of child maltreatment (Dunne et al., 2009). This was assessed using a self-report measure of two items for physical abuse (e.g. "When growing up (before age 18), did your caregiver ever discipline or punish you by physically hitting, spanking, slapping, kicking, or shaking you?") and two items for emotional abuse (e.g. "When growing up (before 18 years) did any person ever discipline or punish you by insulting or criticising you, to make you feel that you were bad, stupid or worthless?"). Items were assessed on a frequency scale of 0 to 3 (0 = Never; 1 = Once or twice; 2 = Three to five times; 3 = More than five times). This instrument was shown to have good internal consistency in the original validation study of seven low, middle, and high-income counties ($0.61 \leq \alpha \leq 0.82$) (Dunne et al., 2009).

Intimate Partner Violence. Caregivers' experience of intimate partner violence over the past month was measured via an adapted version of the Revised Conflict Tactics Scale

Short Form (CTS-2S; 8 items) (Straus et al., 1996). In the Thai study, only respondents who were in a relationship in the previous month were asked to answer this scale, whereas all participants answered this scale in the Philippine study. The scale includes two items on partner negotiation (e.g., "partner suggested a compromise for a disagreement"), five items on physical abuse (e.g., "partner pushed, shoved, or slapped me"), and one item on psychological abuse (e.g., "partner insulted, shouted, yelled, or swore at me"). Responses were coded as a frequency scale of 0 to 3 (0 = Never happened; 1 = Happened once or twice; 2 = Happened three to five times; 3 = Happened more than five times), with higher scores indicating greater levels of intimate partner violence. The scale has been used in LMICs such as South Africa (Ward et al., 2020) and has been found to have good internal consistency ($.68 \leq \alpha \leq .85$) in low-income community samples in the United States (Yun, 2011).

Parent Daily Report on Child Behaviour and Parenting. Day-to-day occurrence of child behaviour problems (34 items) and parenting behaviour (nine items) was assessed via an adapted version of the Parent Daily Report Checklist (PDR) (Patterson et al., 1982). The PDR has been widely used in multiple contexts to assess changes in child and caregiver behaviours during and after programme delivery (Chamberlain et al., 2006; Chamberlain & Reid, 1987). Child behaviour was assessed by asking caregivers whether a child's externalising behaviour (such as hitting others, lying, destructiveness) occurred within the past 24 hours. Parenting behaviour was assessed by asking caregivers about their own behaviour (six items, e.g. whether they shouted or yelled at their child, or praised their child for doing something well) and about parental self-efficacy (three items, e.g. whether they felt that they could not cope with parenting). Item responses were dichotomously scored (0 = No; 1 = Yes) and were summed for each scale on child and parent behaviour. Negative items on the parent subscale were reverse coded so that total scores indicate higher levels of positive parenting behaviour as well as a total PDR parenting score ranging from 0 to 9. The total

score for child behaviour indicates more child behaviour problems within previous 24 hours ranging from 0 to 34. The PDR was administered five times during the RCTs: once at baseline, twice at monthly intervals during the programme, once at post-intervention, and once at follow-up. For the purposes of the current study, only baseline responses were used to predict attendance.

Socio-Economic Factors

In addition to various demographic and health variables, several socio-economic questions were used to assist in assessing living standards and relative poverty beyond direct measures of income. These included household size, employment, household assets, and food consumption.

Household Employment. Household employment was measured using one item from the UNICEF Multiple Indicator Cluster Survey (UNICEF, 2005). This item asked caregivers whether anyone in the household had a job and was scored using a dichotomous variable (0 = No; 1 = Yes).

Household Income Benefits. Information about household income and government benefits was assessed using one item from the UNICEF Multiple Cluster Survey (MICS) (UNICEF, 2005) and one item from the 2015 Thailand National Statistical Office Household Socioeconomic Survey (NSO HHSE) (National Statistical Office and Ministry of Information and Communication Technology, 2016). The MICS item examined the total income of all members of the household, while the NSO HHSE item (only used in Thailand since all participants in the Philippines were recipients of the 4Ps conditional cash transfer system) assessed whether anyone in the household received government benefits.

Household Assets. In Thailand, ownership of household assets was assessed using an item from the UNICEF Multiple Indicator Cluster Survey (UNICEF, 2005). This item allowed for a categorical response concerning collective household ownership of certain

goods, such as cell phone, fridge, car, etc. In the Philippines, household assets were assessed using nine items from the Economic Asset Index which measures relative household poverty based on the Demographic Health Survey (Garcia-Moreno et al., 2006). This scale has successfully been used in over 55 countries, including the Philippines (Montgomery et al., 2000). The tool assessed access to utilities and running water as well as household ownership of consumer goods (cell phone, radio, car etc.).

Household Hunger Level. Assessment of relative poverty according to food insecurity and hunger used five items from the Hunger Scale Questionnaire (Labadarios et al., 2005). Caregivers respond positively or negatively regarding whether they had sufficient money for food, meals were reduced or skipped, or children went to be hungry over the past month, and whether this occurred more than five times. The scale provides intensity scores by summing all the items as well as single occurrence scores.

Primary Outcome

The primary outcome for both RCTs was the frequency of child maltreatment perpetrated by caregivers in the past month. In Thailand, child maltreatment (physical abuse and emotional abuse) was assessed though caregiver self-reports as well as observational measures, as observational coding allows for an independent view of caregiver-child interactions that is less subject to bias (Aspland & Gardner, 2003). In the Philippines, child maltreatment (physical abuse, emotional abuse, and neglect) was only assessed though caregiver self-reports.

Child Maltreatment (Self-Report). An adapted version of the ISPCAN Child Abuse Screening Tool-Trial Caregiver version (ICAST-T) was used to assess the frequency of child maltreatment by caregivers. Twenty items from the Physical Abuse and the Emotional Abuse subscales were included in each country. An additional two items were included to assess caregiver neglect in the Philippines. The ICAST-T measures caregiver reports on the amount

of abuse over the previous month using a frequency score on a scale of 0 to 7, or 8 or more times (e.g., “In the past four weeks, how often did you discipline [name of child] by pushing, grabbing, or kicking him/her?”). Secondary data analysis in this study assesses frequency of child maltreatment for physical abuse (13 items), emotional abuse (seven items), as well as overall frequency of maltreatment in the past month by summing the two subscales (and the two items pertaining to neglect in the Philippines study), with higher scores indicating more use of each type of abuse. This scale was found to have good internal consistency in the original validation study in South Africa ($.77 \leq \alpha \leq .88$) (Meinck et al., 2018), and has successfully been used in studies of PLH interventions in South Africa, Southeastern Europe and the Philippines (Alampay et al., 2018; Frantz et al., 2019; Ward et al., 2020).

Child Maltreatment (Combination of Interview and Observation Assessment). In Thailand, observed maltreatment was also measured through an adapted version of The Home Observation for Measurement of Environment (Home) Inventory Abusive Harsh Parenting subscale (Bradley & Corwyn, 2005). The Home Abusive and Harsh Parenting subscale contains six items that are dichotomously scored. Two of these items were assessed via an interview within home in which data collectors asked the caregiver whether they yelled at the child or used any form of physical punishment in the last week. The remaining four items were assessed via observation and involved data collectors observing whether the caregiver yelled or scolded; used any form of physical restraint; hit, kicked or slapped; and pinched, pulled, or punched pushed the child during the home visit. The scores for this subscale range from 0 to 6 which higher scores indicating more harsh parenting.

Secondary Outcomes

Secondary outcomes assessed include positive parenting; dysfunctional parenting; child monitoring and supervision; child behaviour problems; and parental depression, anxiety, and stress. All measures were caregiver self-reports.

Positive Parenting. Positive parenting practices were assessed via the Parenting Young Children Scale (PARYC) (McEachern et al., 2012). The PARYC measures frequency of positive parenting (seven items), setting limits (seven items), and proactive parenting (seven items) over the previous month on a 7-point Likert-like scale (0 = Never, 6 = Always). The Positive Parenting subscale asks questions such as "Were you able to invite (target child) to play a game with you or share an enjoyable activity?" Setting limits was assessed using items such as "Were you able to set rules on (target child's) behaviour that you were able to enforce?" Items to assess proactive parenting include "Were you able to plan ways to prevent problem behaviour [such as limiting contact with certain peers or not leaving (target child) unsupervised while they are doing certain activities]?" Items of these subscales are summed to form a total positive parenting score (range from 0 to 126) as well as for each subscale (range from 0 to 42) with higher scores indicating higher use of positive parenting behaviours. In the original validation study, the subscales were found to have high construct validity as well as high convergent validity with other similar tests measuring caregiver behaviour (McEachern et al., 2012). The PARYC has previously been used in LMICs, including the Philippines and South Africa (Alampay et al., 2018; Ward et al., 2020).

Dysfunctional Parenting. In Thailand, dysfunctional parenting was assessed using the 10-item Parenting Scale (PS) Over-reactivity subscale which examines caregiver beliefs and attitudes and about authoritarian discipline (Arnold et al., 1993). In the Philippines, the entire 30-item Parenting Scale was used to assess three aspects of dysfunctional discipline practises (laxness, over-reactivity, and verbosity). This instruments measures responses on a 7-point Likert scale in which caregivers are presented with a situation (1 = Most effective; 7 = Most ineffective; i.e., situation: "When I say my child can't do something;" response, score = 1: "I stick to what I said;" or response score = 7: "I let my child do it anyway"). Items sum to create an overall Dysfunctional Parenting score as well as for each of the three subscales,

where higher scores indicate more dysfunctional parenting behaviour. The PS was found to have adequate internal consistency in the original validation study (α = .84) (Arnold et al., 1993) and has also been successfully validated in low-resource settings in the United States (Reitman et al., 2001), and in LMICs such as Panama (Mejia et al., 2015) and Vietnam (Del Vecchio et al., 2017).

Child Monitoring and Supervision. In Thailand, caregiver monitoring and supervision practises were measured using 11 items from an adapted version of the Alabama Parenting Questionnaire (APQ) Monitoring/Supervision subscale (Essau et al., 2006). The APQ assesses parenting practices related to disruptive behaviour in primary school children, but it has also been used in preschool children (Clerkin et al., 2007). The poor monitoring and supervision subscale measures frequency of negative caregiver monitoring and supervision practices over the past month using a 5-point Likert scale (0 = Never; 5 = Always, e.g. "Your child stays out later than he/she is supposed to"). An additional item was included to assess monitoring of social media use (e.g. "Your child uses Facebook, LINE, other social media application, YouTube or the internet without being supervised by an adult"). The items in the subscale are summed to obtain a total score (range from 0 to 55) with higher scores indicating less parental monitoring and supervision. The APQ has been found to have adequate internal consistency (α > .70) (Essau et al., 2006) and has successfully been validated in low-income samples in Australia (Dadds et al., 2003; Elgar et al., 2007).

Child Behaviour Problems. Thirty-six items from the Eyberg Child Behaviour Inventory (ECBI) Intensity Scale and Problem Scale were used to measure child behaviour problems (Eyberg & Ross, 1978). The Intensity Scale measures the frequency of child externalising behaviours in the previous month though a 7-point Liker-like scale (1 = Never occurs; 7 = Always occurs) with scores ranging from 36 to 252. Items include questions such as "How often does (target child) have temper tantrums?". The Problem scale asks caregivers whether

externalising behaviours are currently a problem for them using a dichotomous score (0 = No; 1 = Yes) ranging from 0 to 36. Items on both scales sum to create a total Intensity score and Problem score, with higher Intensity Scale scores indicating greater frequency of problem behaviours and higher Problem Scale sores indicating that caregivers identified more problem behaviours as a problem for them. The ECBI Intensity Scale and Problem Scale have been used extensively in multiple countries to evaluate parenting interventions (Gardner et al., 2006; Gardner et al., 2007; Hutchings, Bywater, Daley, et al., 2007) and were shown to have good internal constancy in the PLH for Young Children trial in South Africa (α = .89, α = .89, respectively) (Ward et al., 2020).

Caregiver Mental Health. Caregiver mental health was assessed using the Depression, Anxiety, and Stress Scale short form (DASS-21) (Lovibond & Lovibond, 2004). This scale assesses the frequency of caregiver Depression, Anxiety, and Stress (seven items per subscale) in the last week using a 4-point Likert scale (0 = Never; 3 = Always). Items include statements such as “I felt that I had nothing to look forward to”. Items sum to create a total scale score as well as for each subscale, with higher scores indicating higher levels of mental health symptoms. Total DASS scores range from 0 to 63, with subscales ranging from 0 to 21. This scale has been shown to have good internal consistency and has successfully been validated in LMICs such as Vietnam showing high consistency (α > .88) (Tran et al., 2013).

Table 1

Outcome Measures for Thailand and the Philippines

Outcome	Measurement	
	Thailand	Philippines
Household and family factors		
Caregiver and child demographics	Multiple Indicator Cluster Survey	Multiple Indicator Cluster Survey
Caregiver physical health	Medical Outcomes Study Short Form – 12 Health Survey	Medical Outcomes Study Short Form – 12 Health Survey
Caregiver disability	Washington Group questions (adapted)	Washington Group questions (adapted)
Caregiver history of abuse during childhood	ISPCAN Child Abuse Screening Tool-Retrospective Version (ICAST-R)	ISPCAN Child Abuse Screening Tool-Retrospective Version (ICAST-R)
Intimate partner violence	Revised Conflict Tactics Scale Short Form (CTS2S): Total intimate partner violence	Revised Conflict Tactics Scale Short Form (CTS2S): Total intimate partner violence
Parent daily report on child behaviour	Parent Daily Report Checklist (PDR): Child problem behaviour	Parent Daily Report Checklist (PDR): Child problem behaviour
Parent daily report on parenting	Parent Daily Report Checklist (PDR): Parenting behaviour	Parent Daily Report Checklist (PDR): Parenting behaviour
Socioeconomic factors		
Household size	Multiple Indicator Cluster Survey	Multiple Indicator Cluster Survey
Household employment	Multiple Indicator Cluster Survey	Multiple Indicator Cluster Survey
Household assets	Multiple Indicator Cluster Survey	Economic Asset Index
Household hunger	The Hunger Scale Questionnaire	The Hunger Scale Questionnaire
Income and benefits	Multiple Indicator Cluster Survey Categories from 2015 National Statistical Office Household Socioeconomic Survey	Multiple Indicator Cluster Survey
Primary outcomes		
Child maltreatment: physical abuse and emotional abuse	1. ISPCAN Child Abuse Screening Tool-Trial Caregiver Version (ICAST-T): Overall maltreatment 1.A ICAST-T: Physical abuse subscale 1.B ICAST-T: Emotional abuse subscale 2. HOME Inventory: Abusive and Harsh Parenting Subscale (interview & observational assessment)	1. ISPCAN Child Abuse Screening Tool-Trial Caregiver Version (ICAST-T): Overall maltreatment 1.A ICAST-T: Physical abuse subscale 1.B ICAST-T: Emotional abuse subscale 1.C ICAST-T: Neglect subscale
Secondary outcomes		
Positive parenting	3. Parenting Young Children Scale (PARYC): Total scale	2. Parenting Young Children Scale (PARYC) – Total scale

Outcome	Measurement	
	Thailand	Philippines
	3.A PARYC Positive parenting subscale	2.A PARYC Positive parenting subscale
	3.B PARYC Setting limits subscale	2.A PARYC Monitoring subscale
	3.C PARYC Proactive parenting subscale	2.C PARYC Proactive parenting subscale
Dysfunctional parenting	4. Parenting Scale (PS): Over-reactivity subscale	3. Parenting Scale (PS): Total scale
		3.A PS Laxness subscale
		3.B PS Over-reactivity subscale
		3.C PSVerbosity subscale
Child monitoring and supervision	5. Alabama Parenting Questionnaire (APQ): Monitoring & supervision subscale	
Child behaviour problems	6.A Eyberg Child Behavior Inventory (ECBI): Problem scale	4.A Eyberg Child Behavior Inventory (ECBI): Problem scale
	6.B Eyberg Child Behavior Inventory (ECBI): Intensity scale	4.B Eyberg Child Behavior Inventory (ECBI): Intensity scale
Caregiver depression, anxiety, and stress	7. Depression, Anxiety, and Stress Scale (DASS): Total scale	5. Depression, Anxiety, and Stress Scale (DASS): Total scale
	7.A DASS Depression subscale	5.A DASS Depression subscale
	7.B DASS Anxiety subscale	5.B DASS Anxiety subscale
	7.C DASS Stress subscale	5.C DASS Stress subscale

Data Analysis

The following section provides a description of the analytic strategies used to answer the two hypotheses of this study. The first section details the steps taken to answer Hypothesis I, including variable selection for the models predicting attendance. The second section provides an extended overview of complier average causal effect (CACE) analysis and how it was used to answer Hypothesis II. Prior to model fitting, scale reliability was also assessed. All analyses were implemented in R version 3.6.1. (R Core Team, 2018).

Scale Reliability

Although the questionnaires were based on validated measures with strong psychometric properties, reliability analyses were conducted to provide evidence that the scores of the items of each scale were reasonably consistent within the population, and that the responses were not simply random noise (Deng & Chan, 2016; McNeish, 2018). Several measures of internal consistency reliability exist, including coefficient alpha (or Cronbach's alpha), coefficient omega, and the greatest lower bound. Among these, coefficient alpha is the most widely used index of scale consistency (Dunn et al., 2014; Hogan et al., 2000). Coefficient alpha is an acceptable estimate of the scale internal consistency under the assumption that items are equally associated to the same underlying construct, also referred to as tau-equivalence (McNeish, 2018). However, when scales are unit-weighted (i.e., the total score is computed by summing the raw scores of individual items) tau-equivalence is frequently violated resulting in an underestimation of the true reliability of a scale (Dunn et al., 2014; Revelle & Zinbarg, 2009; Sijtsma, 2009).

Coefficient omega avoids this assumption and was used as a robust alternative in this study (Dunn et al., 2014; McNeish, 2018; Revelle & Zinbarg, 2009). For positively correlated items, omega total, like coefficient alpha, ranges from zero to one, where zero indicates poor reliability and one perfect internal consistency. As a general guideline, values greater than .70

indicate a reliable scale. The omega function in the R package psych (Revelle, 2018) was used to calculate omega total for the full sample and for the treatment groups individually at each assessment time-point.

Predicting Attendance

The first aim of this study was to determine which factors are associated with attendance in the PLH for Young Children programmes. Generalised linear models, specifically, fixed effects logistic regression models, were used to predict attendance. Logistic regression is a widely used analytic approach to investigate whether a variable has an independent effect on non-normal outcomes (Zhang, 2016). Since the outcome variable (number of sessions attended) was count data, a Poisson error distribution with a logarithm link function was used to approximate the relationship between predictors and attendance (Coxe et al., 2009). This approach transforms the predicted outcome by linearising the nonlinear relationship between the outcome and the predictors using a natural log (Coxe et al., 2009). As a result, the unit of measurement of the predicted scores will be in the natural log of attendance, with the exponents of coefficients equal to the incidence rate ratio (i.e., the relative risk of attending/missing a session).

Many potential predictors of attendance were identified in the literature review. Since testing numerous predictor variables individually may increase the risk of falsely concluding that a predictor has a significant effect (Type I error), and including many predictors concurrently may increase the risk of falsely concluding that a predictor does not have a significant effect (Type II error) (Bursac et al., 2008; Zhang, 2016), a model building strategy was developed to minimize the risk of Type I and II errors.

The common approach to logistic model building is minimisation of predictor variables while ensuring that these still reflect the true outcomes of the data (the most parsimonious model) (Bursac et al., 2008; Zhang, 2016). There are many strategies to achieve

this, including variable selection algorithms such as ridge and lasso regression. While there is no one correct method, mechanical approaches such as these may carry a host of limitations, especially for smaller sample sizes (Bursac et al., 2008). Thus, a purposeful variable selection approach, informed by the literature review on the factors associated with attendance in parenting interventions, was used. Purposeful variable selection has been widely used in exploratory analyses and draws on a range of steps proposed by Hosmer and Lemeshow (2000).

The details of the purposeful model building process used were as follows: first, possible predictors were grouped conceptually into the four domains identified in the literature review: 1) socio-economic and educational resources and barriers, 2) social and health recourses and barriers, 3) caregiver and child behaviour, and 4) sociodemographic characteristics. These sets of predictors were primarily used to ensure that each domain identified was represented in the model building process. Next, univariable analyses were conducted to explore the unadjusted association between an individual predictor and attendance (Hosmer & Lemeshow, 2000). Any predictor variable with association of $p \leq .25$ was then entered into a multivariable model. This cut-off, widely supported by literature, was used because more traditional cut-off points of $p < .05$ can sometimes fail to identify predictors known to be important (Bendel & Afifi, 1977; Hosmer & Lemeshow, 2000; Mickey & Greenland, 1989).

In addition, a parenting group assignment variable was included as fixed effects in the multivariable model because participants were nested within parenting groups. With only four parenting groups ($n = 15$ in each group) both in Thailand and the Philippines, there was an insufficient number of groups and participants within each group to accurately test whether participants' attendance significantly differed between parenting group.

Consequently, a parenting group variable was only included to account for dependence in the data and not to test the influence of parenting groups on participant attendance.

Spearman's correlation coefficients were examined to reveal any strong relationships between predictor variables selected for the multivariable model. Variables with strong inter-correlations (<.60) were removed one at a time to control for multicollinearity, and new smaller models was fitted each time (Zhang, 2016). These models where then compared using partial likelihood ratio tests and Akaike Information Criterion values to identify the most parsimonious fit.

Due to the small sample size (only the intervention arm was included in the logistic regression analyses to examine associations with attendance), all predictor variables were tested as main effects rather than also testing for interactions amongst predictors. Once the model with the most parsimony was identified, continuous variables were checked for their linearity in relation to the logarithm of attendance (Zhang, 2016). This was done by visually inspecting the scatter plot between each predictor and the logarithm values of attendance.

Finally, the fit of the multivariable model was assessed using the summary measures of goodness of fit and the regression diagnostics. The former was assessed using the Hosmer-Lemeshow test (Hosmer & Lemeshow, 2000) which measures the difference between the observed and the fitted vales of the model. To ensure that assumptions were upheld and that the model fits across the entire range of retained covariates, regression diagnostics were assessed, including an inspection of plots of residuals against fitted values and predictors, influential values in the continuous predictors, and multicollinearity among predictors (Zhang, 2016).

In line with emerging guidelines against solely reporting on statistical significance, as this can lead to misinterpretation of findings (e.g., Nickerson, 2000; Schmidt, 1996; Ziliak & McCloskey, 2008), effect sizes are reported as a means of assessing the strength of

association between predictor and outcome along with confidence intervals (CIs) as an indication of the preciseness of the effect. Since logistic regression models the impact of predictors on the risk of the event – for this study, the risk of missing a session - effect sizes are reported using incidence rate ratios (IRRs) along with their 95% CIs. IRRs are interpreted as follows: an IRR equal or close to 1.00 suggests that there is little or no difference in risk of not attending (i.e. the incidence of attendance is the same in each group). An IRR greater than 1.00 suggests an increased chance of attending, whereas an IRR smaller than 1.00 suggests a reduced risk of attending. For example, an IRR of 0.75 would indicate that there is a 0.75 (or 25%) times lower likelihood of attending compared to the comparison group, whereas an IRR of 1.25 corresponds to a 1.25 (or 25%) times greater likelihood of attending. In addition, robust standard errors were used for the parameter estimates in the multivariable models to control for minor violations of the Poisson distribution assumption that the variance of the outcome equals its mean (Cameron & Miller, 2015). The R package sandwich (Zeileis, 2006) was used to obtain parameter estimates and their robust standard errors along with *p*-values calculated accordingly.

Complier Average Causal Effects (CACE)

The second aim of this study was to estimate the effects of PLH for Young Children on caregivers and children in Thailand and the Philippines while also accounting for attendance in the intervention. This section outlines the key analytic steps that were taken to achieve this aim, including defining compliance in the intervention group, providing a detailed overview of CACE analysis and its assumptions, including covariates in the CACE models, and supplementing CACE estimates with intention-to-treat (ITT) estimates.

Defining Compliance. Defining compliance, or engagement, is an important first step in CACE analysis. Following recommendations by Angrist et al. (1996), a dichotomous indicator of engagement (0 = Not engaged; 1 = Engaged) is required for the identification of

CACE estimates. Although full engagement or no engagement may be obvious in some studies, determining an "all or none" cut-off of engagement is challenging in parenting programmes. For example, if engagement is defined as attending the total number of intervention sessions, the assumption is that caregivers who attended fewer sessions were not affected by the programme. However, setting a low threshold of engagement may result in large variations in the degree to which participants attended the programme, while a more stringent threshold decreases the sample size among engagers thereby also decreasing the quality CACE estimates (Berg et al., 2017; Stuart et al., 2008). Given this dilemma and the lack of prior research regarding the exact dosage required to improve caregiver and child outcomes in parenting programmes, sensitivity analyses were conducted using two different definitions of engagement: 1) caregivers who attended at least 50% of the total number of sessions were classified as moderate attenders; 2) caregivers who attended at least 75% of the programme were deemed high attenders, with separate models addressing each definition. These definitions of engagement are consistent with behavioural and educational studies that have examined CACE impacts (Ashworth et al., 2020; Berg et al., 2017; Dishion et al., 2014).

Using the less stringent definition of engagement, 54 (90%) of the caregivers in the intervention group in Thailand were considered moderate attenders while the remainder were classified "not engaged". The "dosage" received by moderate attenders ranged from 4 to 8 sessions (M = 7.20, SD = 1.14). In the Philippines, 43 (71.7%) of the caregivers in the intervention group were defined as moderate attenders, and their "dosage" received ranged between 6 and 12 sessions (M = 9.16, SD = 1.86). Using more stringent criterion, 47 (78.3%) caregivers in Thailand were defined as high attenders and the rest as not engaged. The "dosage" received by high attending caregivers ranged from 6 to 8 sessions (M = 7.57, SD =

0.62). In the Philippines, 26 (43.3%) caregivers were classified as having high attendance, with a dosage ranging between 9 and 12 sessions (M = 10.42, SD = 1.10).

Overview of CACE. Assuming binary treatment assignment (0 = Control group; 1 = Intervention group) and binary treatment received (e.g., 0 = Did not attend 50% or 75% of session; 1 = Attended at least 50% or 75% of sessions) for each individual, CACE methods define four possible participant behaviours namely: compliers, never-takers, defiers, and always-takers (Angrist et al., 1996; Little & Yau, 1998). In the case of a two-armed parenting intervention, *compliers* are caregivers who receive the programme (i.e., attend programme sessions) if they are assigned to the intervention group or who do not receive/attend the programme if they are in the control group. *Never-takers* are caregivers who do not attend the programme, regardless of the group they were assigned to. In contrast, *defiers* are caregivers who do not attend the programme when they are in the intervention group but who do attend when they are in the control group. Lastly, *always-takers* are caregivers who, regardless of their treatment assignment, always attend the programme. Typically, prevention interventions such as parenting programmes only allow participants in the intervention group access to programme sessions. Therefore, since the control group cannot access the intervention, always-takers and defiers are assumed to be absent (Connell, 2009; Stuart et al., 2008). There are thus only two compliance types that need to be accounted for (compliers and never-takers), resulting in a binary indicator of compliance. This indicator allows for an estimation of programme effects based on outcomes for compliers in the intervention group compared to outcomes for compliers in the control group who would have complied in the programme had they been randomised to receive it (Connell, 2009).

Validating CACE Assumptions. Causal interpretation with CACE is possible under five key assumptions (Angrist et al., 1996). The following section will describe these assumptions and discuss the degree to which they were met in the context of this study.

1. *Random assignment*. CACE analysis requires that participation in the study was randomly assigned. Since both studies involved caregivers being randomly assigned to either the intervention (PLH for Young Children) or the control group, this assumption is satisfied.
2. *Stable Unit Treatment Value Assumption.* This assumption requires that the outcomes for each participant are not influenced by the outcomes of other participants. In the context of a group-based intervention such as PLH, this assumption is potentially problematic because the composition of the group may influence caregivers' outcomes. However, there has been little progress in developing methods to account for interactions between participants in both ITT and CACE approaches (Stuart et al., 2008). Therefore, this study will assume that this assumption holds well enough, although it is noted as a potential shortcoming of the analyses.
3. *No defiers.* This assumption, also referred to as "monotonicity", implies that being in the intervention group can only increase engagement and not decrease it (i.e. there are no defiers). As mentioned above, defiers are participants who do not comply with the treatment assignment they were randomised to. In this study, none of the caregivers assigned to the control group received the programme. It is therefore safe to assume that there were no defiers in the control group. Given that caregivers did not have access to the programme unless they were randomly assigned to it, the same lack of defiers should also be found in the intervention group. That is, always-taker membership is zero due to the design of the study. This assumption is necessary as it helps identify compliers by reducing the number of compliance types for whom estimates are calculated, thereby permitting a binary compliance variable (Connell, 2009).

4. *Rate of compliance is not zero.* This assumption requires that the compliance, or engagement, rate in the intervention group is not zero. In other words, there must be some caregivers who attend sessions. As shown in the description of engagement above, there is evidence that there were some caregivers who attended the programme sessions. Thus, this assumption has been met.
5. *Exclusion restriction*. This assumption states that there are no intervention effects for noncompliers, i.e., in order to benefit from the intervention caregivers must actually participate. Since engagement is defined as attending at least 50% or 75% of the total number of programme sessions, this assumption implies that caregivers who attended below the engagement cut-off received no benefits. Therefore, the exclusion restriction is somewhat questionable for this study because noncompliers in the intervention group may have participated in some of the group sessions or the home visits or received phone calls and text messages. Violations of this assumption have been shown to lead to biased CACE estimates (Jo, 2002). However, including baseline covariates to predict engagement can reduce bias and result in accurate estimates (Frangakis & Rubin, 1999; Jo, 2002).

Estimating Programme Effects Using CACE

Prior to modelling, descriptive statistics of each outcome were examined along with distribution plots to determine the appropriate regression type. In Thailand, the ICAST total scale as well as the physical and emotional subscales, the ECBI child problem behaviour scale, and the DASS total scale were deemed Poisson distributions (see Appendix G). In the Philippines, only the ICAST total scale, the physical abuse and emotional abuse subscales, and the ECBI child problem behaviour had Poisson-like distributions (see Appendix H). The remainder of the scales were deemed normal enough for linear modelling. Robust standard errors were used to account for any non-normality in the data (Dunn et al., 2005).

To obtain CACE estimates, this study used the two-stage least squares approach. In the first stage, attendance was regressed on the instrumental variable (treatment allocation/randomisation), and a predicted value for attendance was obtained for each participant in the intervention and the control group. In the second stage, the outcome (for example, child maltreatment) was regressed on these predictions of attendance, and the slope of the second-stage regression was used as an estimate of the causal programme effect.

For normally distributed outcomes, this structural model was fit using the ivreg function from the R package AER (Kleiber & Zeileis, 2008) which computes these two stages jointly. Since current implementation of two-stage least squares approaches in R does not support non-linear models (Sjolander & Martinussen, 2019), all non-normally distributed outcomes were computed manually. This was done using the lm function in the first stage to obtain predictions of attendance for each participant. In the second stage, a logistic regression model with a Poisson error distribution and a logarithm link was used to obtain point estimates via the glm function.

Estimation of CACE effects commonly only considers information from a single time point (Wu, 2016). Since the present study assessed outcomes both at post-intervention and at follow-up, an analysis of covariance approach was used where the baseline measurement was treated as a covariate for measurement at post-test and follow-up, as well as a predictor for compliance status (Wu, 2016). That is, two separate models were built, one predicting post-intervention effects while controlling for baseline scores, and the second predicting follow-up effects while controlling for baseline scores.

Furthermore, two CACE models were also built for each outcome at each time-point, using the moderate and highly engaged definitions of compliance described above. As proposed by Angrist and colleagues (1996), engagement status was treated as a dichotomous variable (0 = Non-engagers; 1 = Engagers). Since the control group did not have access to

programme sessions, all participants in the control group were also coded as non-engagers. In these models, it was assumed that non-engagers received no benefits from the programme (referred to the exclusion restriction assumption). However, given the possibility that the exclusion restriction was violated, a third CACE model was built which included baseline predictors of attendance as covariates (Jo, 2002; Jo et al., 2008). Variables included in this model were identified in the analyses of predictors of attendance, and thus differed by country. Child age and child gender were also included as covariates in analyses to control for potential dependency of these variables on outcomes.

Given the small amounts of missing data (only 6 cases were lost to follow-up in each RCT), an imputation approach such as multiple imputation was not deemed necessary (Schulz & Grimes, 2002). However, there is minimal research which provides guidelines for handling missing data within an instrumental variable regression framework in R. Therefore, this study ignored the potential biases introduced by dropping cases with missing data and employed listwise deletion, i.e., the analysis included all eligible randomised caregivers who provided post-intervention and follow-up data. This meant that in the models predicting post-intervention outcomes, four participants (3.3% in each study) were excluded due to dropout or absenteeism on the day of assessment, resulting in a sample size of 116 caregivers in each RCTs. In the models predicting programme effects at follow-up, an additional two caregivers (1.7% in each study) were dropped due to attrition, resulting in a sample size of 114 caregivers in each RCT. While this approach is common in studies utilising CACE analysis (e.g. Knox et al., 2014; Yau & Little, 2001), it is noted as a limitation of this study.

To help interpret results, estimates for the CACE models were compared to intention-to-treat (ITT) estimates in which participants were analysed according to their allocation assignment (intervention versus control) regardless of attendance or dropout using restricted maximum likelihood estimation. As with the CACE models, treatment assignment and

baseline measures of the outcome were regressed on the outcome at post-test and at follow-up. Similarly, logistic regression analyses were employed for normally distributed outcomes whereas linear regression analyses were used for non-normally distributed outcomes. CACE models were compared to ITT models descriptively using IRRs for the logistic regression models and unstandardized coefficients for the linear models. All participants in all the models were analysed according to their treatment allocation.

To conclude, this chapter described the PLH for Young Children programme implementation in Thailand and the Philippines, including the study deigns, and procedures. This section also provided an overview of the measures selected for the secondary data analyses and then proceeded with a discussion of the specific analysis strategy used to implement the secondary data analysis. The next chapter will present the results of the secondary data analyses.

Chapter 3: Results

Scale Reliability

Table 2 and Table 3 present the results of the reliability analyses for PLH for Young Children Thailand and PLH for Young Children Philippines, respectively. All but two of the scales in the Thailand study had acceptable reliability, with a coefficient omega value greater than .70 at one or more of the assessment time points. The PDR parent scale (which measures the occurrence of particular parenting behaviours) and the CTS-2S scale (which measures the frequency or count of exposure to intimate partner violence) both had a coefficient omega value smaller than .70 at all the assessment time-points for either the total sample for one or for allocation group. Since reliability statistics such as coefficient omega assess how inter-related items on a scale are, they are not always suitable for measures which count the occurrence of a behaviour (Weissinger et al., 1992). Rather, this statistic is more suited to measuring latent constructs such as parenting stress or depression. For this reason, these two scales were retained in the outcome analyses, despite their poor reliability estimates. Moreover, at least half of the items on the CTS-2S scale were removed (for reliability analyses only) for the intervention group due to zero variance in the items - resulting in four or fewer items. Thus, poor reliability may likely also have been due to loss of items since fewer items typically have lower reliability values (Cortina, 1993). For the Philippines study, all scales, except for the Verbosity subscale, had a coefficient omega of greater than .70 at one or more of the assessment time points. Since the Verbosity subscale measures a latent construct of dysfunctional parenting, it was not retained in the outcome analyses.

Table 2

Scale Reliability at Baseline, Post-test, and Three Months Follow-up: Thailand

Measure	Assessment time-point					
	Baseline		One-month post-test		Three-month follow-up	
	Omega	No. items dropped[1]	Omega	No. items dropped[1]	Omega	No. items dropped[1]
ICAST Total						
Full sample	.92	1	.90	2	.89	2
Control	.94	1	.90	2	.92	3
Intervention	.86	6	.88	7	.72	8
ICAST (Physical abuse)						
Full sample	.92	1	.87	2	.83	2
Control	.89	1	.89	2	.87	3
Intervention	.86	5	.86	2	.89	7
ICAST (Emotional abuse)						
Full sample	.90	0	.85	0	.83	0
Control	.93	0	.87	0	.86	0
Intervention	.84	1	.80	2	.62	1
PARYC Overall						
Full sample	.85	0	.93	0	.94	0
Control	.90	0	.94	0	.94	0
Intervention	.87	0	.90	0	.93	0
PARYC (Positive parenting)						
Full sample	.82	0	.80	0	.84	0
Control	.77	0	.82	0	.77	0
Intervention	.84	0	.82	0	.85	0
PARYC (Setting limits)						
Full sample	.87	0	.91	0	.91	0
Control	.88	0	.88	0	.90	0
Intervention	.88	0	.84	0	.86	0
PARYC (Proactive parenting)						
Full sample	.83	0	.86	0	.89	0
Control	.89	0	.90	0	.88	0
Intervention	.84	0	.88	0	.90	0
PS (Over-reactivity)						
Full sample	.83	0	.83	0	.81	0
Control	.84	0	.85	0	.86	0
Intervention	.80	0	.75	0	.74	0
APQ (Poor monitoring)						
Full sample	.78	0	.86	0	.88	0
Control	.85	0	.89	0	.90	0
Intervention	.70	1	.77	0	.84	1
DASS Total						
Full sample	.94	0	.93	0	.95	0
Control	.95	0	.92	0	.96	0
Intervention	.93	0	.93	0	.93	0
DASS (Depression)						
Full sample	.85	0	.83	0	.86	0
Control	.86	0	.81	0	.88	0
Intervention	.81	0	.88	0	.87	0
DASS (Anxiety)						
Full sample	.90	0	.82	0	.93	0
Control	.93	0	.87	0	.93	0
Intervention	.89	0	.87	0	.92	0
DASS (Stress)						

Measure	Assessment time-point					
	Baseline		One-month post-test		Three-month follow-up	
	Omega	No. items dropped[1]	Omega	No. items dropped[1]	Omega	No. items dropped[1]
Full sample	.88	0	.83	0	.90	0
Control	.90	0	.83	0	.93	0
Intervention	.85	0	.84	0	.88	0
ECBI (Intensity Scale)						
Full sample	.94	0	.96	0	.96	0
Control	.95	0	.96	0	.97	0
Intervention	.94	0	.96	0	.95	0
ECBI (Problem Scale)						
Full sample	.95	0	.97	0	.97	1
Control	.95	0	.97	0	.97	1
Intervention	.96	0	.97	1	.95	2
PDR (Child behaviour)						
Full sample	.90	0	.92	0	.93	0
Control	.91	0	.92	0	.94	0
Intervention	.89	0	.92	2	.91	6
PDR (Parent behaviour)						
Full sample	.63	0	.61	0	.62	0
Control	.62	0	.58	0	.63	0
Intervention	.70	0	.70	0	.72	1
CTS-2S						
Full sample	.89	0	.88	0	.96	0
Control	.91	0	.93	0	.97	0
Intervention	.40	4	.51	5	.08	4

Note. [1]Items with zero variance were only removed to compute omega but not for the outcome analyses.

ICAST = International Society for the Prevention of Child Abuse and Neglect Child Abuse Screening Tool; PARYC = Parenting Young Children Scale; PS = Parenting Scale; APQ = Alabama Parenting Questionnaire; DASS = Depression, Anxiety, and Stress Scale; ECBI = Eyberg Child Behaviour Inventory; PDR = Parent Daily Report Checklist; CTS-2S = Revised Conflict Tactics Scale Short Form.

Table 3

Scale Reliability at Baseline, Post-test, and Six Months Follow-up: Philippines

Measure	Assessment time-point					
	Baseline		Immediate post-test		12-months follow-up	
	Omega	No. items dropped[1]	Omega	No. items dropped[1]	Omega	No. items dropped[1]
ICAST Total						
Full sample	.92	0	.93	0	.92	0
Control	.95	0	.95	0	.94	0
Intervention	.90	0	.63	0	.89	0
ICAST (Physical abuse)						
Full sample	.90	0	.90	0	.83	0
Control	.91	0	.93	0	.84	0
Intervention	.90	0	.73	0	.83	0
ICAST (Emotional abuse)						
Full sample	.85	0	.85	0	.88	0
Control	.90	0	.89	0	.90	0
Intervention	.81	0	.55	0	.62	0
PARYC Overall						
Full sample	.86	0	.93	0	.91	0
Control	.90	0	.91	0	.93	0
Intervention	.83	0	.94	0	.91	0
PARYC (Positive parenting)						
Full sample	.78	0	.88	0	.71	0
Control	.79	0	.87	0	.80	0
Intervention	.83	0	.86	0	.78	0
PARYC (Monitoring)						
Full sample	.80	0	.86	0	.88	0
Control	.90	0	.91	0	.93	0
Intervention	.76	0	.88	0	.86	0
PARYC (Proactive parenting)						
Full sample	.81	0	.89	0	.88	0
Control	.85	0	.83	0	.89	0
Intervention	.79	0	.93	0	.88	0
PS Total						
Full sample	.75	0	.75	0	.79	0
Control	.82	0	.77	0	.76	0
Intervention	.73	0	.66	0	.83	0
PS (Laxness)						
Full sample	.70	0	.66	0	.76	0
Control	.78	0	.58	0	.80	0
Intervention	.52	0	.71	0	.72	0
PS (Over-reactivity)						
Full sample	.73	0	.73	0	.79	0
Control	.78	0	.72	0	.76	0
Intervention	.67	0	.72	0	.86	0
PS (Verbosity)						
Full sample	.52	0	.53	0	.65	0
Control	.61	0	.63	0	.58	0
Intervention	.55	0	.62	0	.69	0
DASS Total						
Full sample	.91	0	.92	0	.91	0
Control	.92	0	.93	0	.91	0
Intervention	.88	0	.93	0	.93	0
DASS (Depression)						

Measure	Assessment time-point					
	Baseline		Immediate post-test		12-months follow-up	
	Omega	No. items dropped[1]	Omega	No. items dropped[1]	Omega	No. items dropped[1]
Full sample	.77	0	.83	0	.82	0
Control	.80	0	.84	0	.78	0
Intervention	.73	0	.80	0	.81	0
DASS (Anxiety)						
Full sample	.83	0	.85	0	.86	0
Control	.81	0	.89	0	.85	0
Intervention	.86	0	.83	0	.89	0
DASS (Stress)						
Full sample	.86	0	.86	0	.82	0
Control	.89	0	.89	0	.82	0
Intervention	.82	0	.84	0	.86	0
ECBI (Intensity Scale)						
Full sample	.90	0	.95	0	.93	0
Control	.91	0	.96	0	.93	0
Intervention	.90	0	.93	0	.93	0
ECBI (Problem Scale)						
Full sample	.93	0	.95	0	.96	0
Control	.95	0	.96	0	.97	0
Intervention	.91	1	.93	0	.95	1
PDR (Child behaviour)						
Full sample	.83	0	.89	2	.86	1
Control	.87	0	.90	3	.85	2
Intervention	.75	1	.89	2	.89	3
PDR (Parent behaviour)						
Full sample	.59	0	.48	0	.70	0
Control	.60	0	.57	0	.69	0
Intervention	.53	0	.76	0	.81	0
CTS-2S						
Full sample	.80	0	.96	0	.93	0
Control	.88	0	.96	0	.96	0
Intervention	.74	0	.95	0	.82	1

Note. [1]Items with zero variance were only removed to compute omega but not for outcome analyses. ICAST = International Society for the Prevention of Child Abuse and Neglect Child Abuse Screening Tool; PARYC = Parenting Young Children Scale; PS = Parenting Scale; APQ = Alabama Parenting Questionnaire; DASS = Depression, Anxiety, and Stress Scale; ECBI = Eyberg Child Behaviour Inventory; PDR = Parent Daily Report Checklist; CTS-2S = Revised Conflict Tactics Scale Short Form.

Baseline Sample Characteristics of Families

Thailand

Sample characteristics at baseline are summarised in Table 4. Children of caregivers had a mean age of 5.23 years (*SD* = 1.91), and 47 (39.5%) were female. The mean age of caregivers was 43.84 years (*SD* = 13.56). Almost all the caregivers were female (*n* = 116, 96.7%) and almost half were grandparents (*n* = 57, 47.5%). Three-quarters of caregivers were married (*n* = 90, 75%) and one-third reported fair health (*n* = 40, 33.3%). Over a half of the sample had not completed high school (*n* = 75, 62.5%) and 50 (41.7%) were unemployed. Almost all caregivers reported a household income source from wages and salaries (*n* = 108, 90%), yet 33.3% (*n* = 40) reported having run out of money to buy food in the last 30 days. There were no statistically significant differences between allocation groups expect for caregiver gender, with all four male caregivers allocated to the intervention group (*t* = -2.06, *p* = 0.045).

Philippines

A description of baseline sample characteristics is reported in Table 5. Children of caregivers had a mean age of 3.81 years (*SD* = 1.25), and just over half (*n* = 64, 53.3%) were female. The mean age of caregivers was 36.11 years (*SD* = 6.56) and all were female (*n* = 120, 100%). Almost all caregivers reported being the children's biological mothers (*n* = 116, 96.7%) – the rest were their grandmothers (*n* = 4, 3.3%). Roughly half of the caregivers were married (*n* = 61, 50.8%) and were in fair health (*n* = 67, 55.8%). Almost half of the sample had not completed high school (*n* = 59, 49.2%). While almost all caregivers reported a household income source from wages and/or salaries (*n* = 101, 84%), approximately one-third were unemployed (*n* = 42, 35%). Just over a half of caregivers reported running out of money for food in the last 30 days (*n* = 62, 51.7%). There were no statistically significant

differences between allocation groups except for that the control group reported higher rates of running out of money for food than the intervention group ($t = 2.21$, $p = 0.028$).

Table 4

Family Characteristics at Baseline: Thailand

Variable	Total ($N = 120$)	Control ($n = 60$)	Intervention ($n = 60$)	Group comparison[1] *p*
Caregiver age, *M* (*SD)*	43.84 (13.56)	42 (13.62)	45.68 (13.37)	0.138
Caregiver female, *n* (%)	116 (96.7)	60 (100)	56 (93.3)	0.045
Child age, *M* (*SD*)	5.23 (1.91)	5.03 (1.85)	5.40 (2)	0.296
Child female, *n* (%)	47 (39.5)	25 (41.7)	22 (36.7)	0.579
Relationship to child				0.266
Biological parent, *n* (%)	56 (46.7)	30 (50)	26 (43.3)	
Stepparent, *n* (%)	1 (0.8)	1 (1.7)		
Grandparent, *n* (%)	57 (47.5)	27 (45)	33 (55)	
Great-grandparent, *n* (%)	5 (4.2)	2 (3.3)	3 (5)	
Cousin, *n* (%)	1 (0.8)		1 (1.7)	
Marital status				0.197
Unmarried, *n* (%)	6 (5)	3 (5)	3 (5)	
Married, *n* (%)	90 (75)	43 (71.7)	47 (78.3)	
Separated, *n* (%)	12 (10)	6 (10)	6 (10)	
Widowed, *n* (%)	12 (10)	8 (13.3)	4 (6.7)	
Caregiver health[2]				0.484
Excellent, *n* (%)	3 (2.5)	1 (1.7)	2 (3.3)	
Very good, *n* (%)	18 (15)	12 (20)	4 (6.7)	
Good, *n* (%)	37 (30.8)	14 (23.3)	23 (38.3)	
Fair, *n* (%)	40 (33.3)	18 (30)	22 (36.7)	
Poor, *n* (%)	22 (18.3)	15 (25)	7 (11.7)	
Caregiver education				0.181
Not completed primary school, *n* (%)	3 (2.5)	1 (1.7)	2 (3.3)	
Completed primary school, *n* (%)	34 (28.3)	16 (26.7)	17 (28.3)	
Some secondary school, *n* (%)	36 (30)	17 (28.3)	19 (31.7)	
Completed secondary school, *n* (%)	27 (22.5)	14 (23.3)	13 (21.7)	
Post-high school, *n* (%)	17 (14.2)	10 (16.7)	7 (11.7)	
Caregiver unemployed, *n* (%)	50 (41.7)	26 (43.3)	24 (40)	0.714
Income source				
No household income, *n* (%)	2 (1.7)		2 (3.3)	0.411
Wages, salaries, or profits, *n* (%)	108 (90)	55 (91.7)	53 (88.3)	
Remittances, *n* (%)	4 (3.3)	1(1.7)	3 (5)	
Other incomes, *n* (%)	6 (5)	4 (6.7)	2 (3.3)	
Ran out of money to buy food in the last 30 days, *n* (%)	40 (33.3)	19 (31.7)	21 (35)	0.701

Note. [1]t-tests were conducted to compare intervention and control group scores. [2] Medical Outcomes Study Short Form – 12 Health Survey, range = 1-5.

Table 5

Family Characteristics at Baseline: Philippines

Variable	Total (N = 120)	Control (n = 60)	Intervention (n = 60)	Group comparison[1] p
Caregiver age, *M* (*SD)*	36.11 (6.56)	36.6 (6.81)	35.62 (6.32)	0.413
Caregiver female, *n* (%) [1]	120 (100)	60 (100)	60 (100)	
Child age, *M* (*SD*)	3.81 (1.25)	3.80 (1.22)	3.82 (1.30)	0.942
Child female, *n* (%)	64 (53.3)	33 (55)	31 (51.7)	0.717
Relationship to child				1.000
Biological parent, *n* (%)	116 (96.7)	58 (96.7)	58 (96.7)	
Grandparent, *n* (%)	4 (3.3)	2 (3.3)	2 (3.3)	
Marital status				0.143
Single, *n* (%)	4 (3.3)	2 (3.3)	2 (3.3)	
Unmarried but in a relationship, *n* (%)	53 (44.2)	23 (38.3)	30 (50)	
Married, *n* (%)	61 (50.8)	33 (55)	28 (46.7)	
Separated, *n* (%)	1 (0.8)	1 (1.7)		
Widowed, *n* (%)	1 (0.8)	1 (1.7)		
Caregiver health [2]				0.126
Excellent, *n* (%)	2 (1.7)	1 (1.7)	1 (1.7)	
Very good, *n* (%)	8 (6.7)	5 (8.3)	3 (5)	
Good, *n* (%)	36 (30)	23 (38.3)	13 (21.7)	
Fair, *n* (%)	67 (55.8)	27 (45)	40 (66.7)	
Poor, *n* (%)	7 (5.8)	4 (6.7)	3 (5)	
Caregiver education				0.173
Not completed primary school, *n* (%)	13 (10.8)	4 (6.7)	9 (15)	
Completed primary school	10 (8.3)	6 (10)	4 (6.7)	
Some secondary/ high school, *n* (%)	36 (30)	17 (28.3)	19 (31.7)	
Vocational school, *n* (%)	5 (4.2)	3 (5)	2 (3.3)	
Completed secondary/ high school, *n* (%)	38 (31.7)	19 (31.7)	19 (31.7)	
Post-high school, *n* (%)	18 (15)	11 (18.3)	7 (11.7)	
Caregiver unemployed, *n* (%)	42 (35)	23 (38)	19 (31.7)	0.448
Income source				0.321
No household income, *n* (%)	3 (2.5)	1 (1.7)	2 (3.3)	
Money earned from selling things informally, *n* (%)	14 (11.7)	9 (15)	5 (8.3)	
Salaried and/or wages, *n* (%)	101 (84.2)	48 (80)	53 (88.3)	
Remittances, *n* (%)	1 (0.8)	1 (1.7)		
Other incomes, *n* (%)	1 (0.8)	1 (1.7)		
Ran out of money to buy food in the last 30 days, *n* (%)	62 (51.7)	37 (61.7)	25 (41.7)	0.028

Note. [1]t-tests were conducted to compare intervention and control group scores. [2]All caregivers were female thus no t-test was necessary for this variable. [2] Medical Outcomes Study Short Form – 12 Health Survey, range = 1-5.

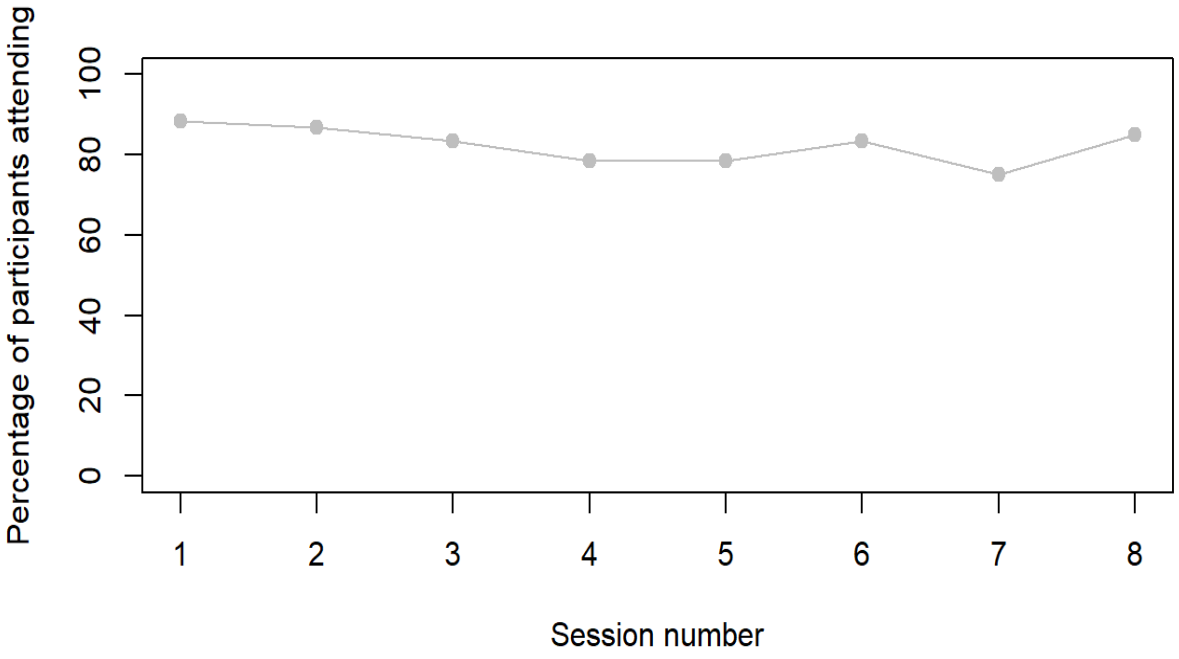
Programme Attendance

Thailand

The average overall attendance rate, which included both attendance at group sessions and home visits, was 6.58 (*SD* = 2.20) of 8 sessions, or 82.3% of the intervention. Of the 60 caregivers who were enrolled in the intervention, 54 (90%) attended at least half (four group sessions or home visits) of the programme. Forty-seven (78%) participants attended six or more sessions and 30 (50%) attended all 8 sessions (i.e. 100% of the programme). Figure 3 plots the overall attendance across the eight intervention sessions. Overall, attendance was relatively consistent over the eight sessions. The highest attendance was at session 1 (88.3%), while lowest was at session 7 (75%). Attendance at session 8 was also very high (85%) – perhaps because participants knew that this was a celebratory session which included receiving a certificate of completion.

Figure 3

Overall Attendance Across Eight Sessions in Thailand

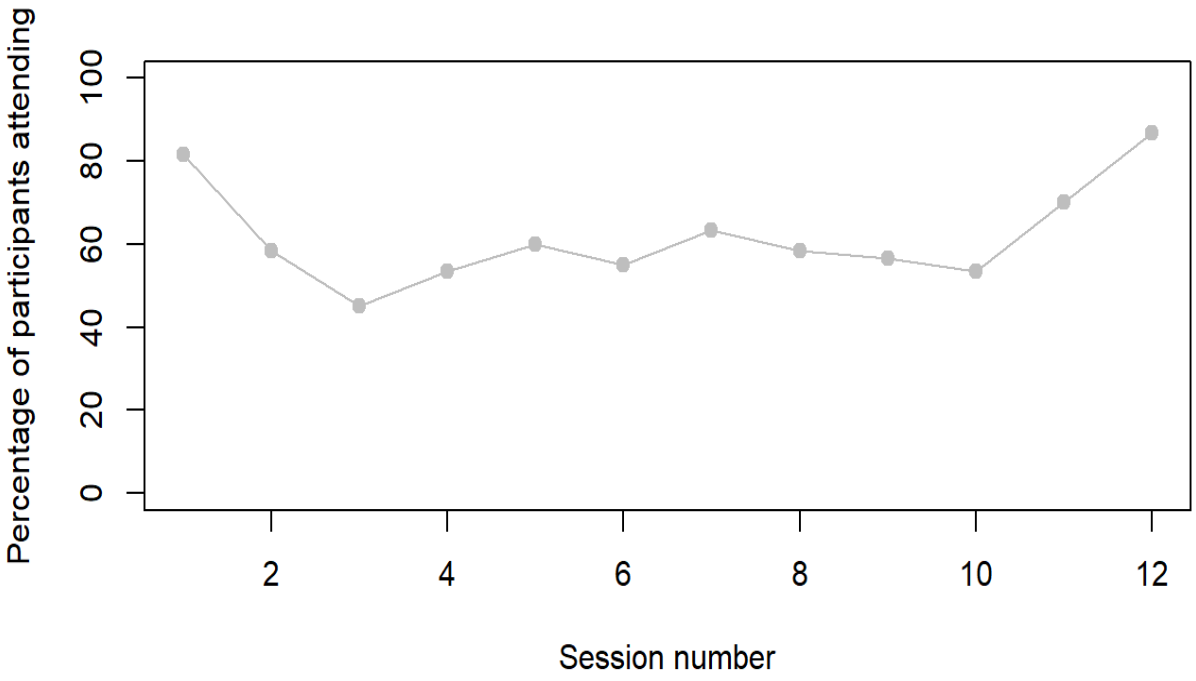


Philippines

The average overall attendance rate was 7.42 (*SD* = 3.35) of 12 sessions, or 61.8% of the programme. Forty-three (71.67%) caregivers attended half or more (six group sessions or home visits) of the programme, 26 (43.3%) caregivers attended nine or more sessions, while only 5 (8.3%) attended all 12 sessions. Figure 4 plots overall attendance across the 12 intervention sessions. There were considerable fluctuations across sessions, with a dip of approximately 20% between session 1 and session 3. However, attendance gradually increased from session 4 onwards. These fluctuations may be attributed to logistical barriers including the distance caregivers had to travel to and from the venues. This was an especially important factor for non-attendance during extreme weather that resulted in heavy rainfall and flooding in the community.

Figure 4

Overall Attendance Across Twelve Sessions in the Philippines



Predicting Attendance and its Impacts on Outcomes

The main analyses proceeded in two stages in order to address the two study hypotheses. In the first stage, logistic regression analyses were utilised to model the relationship between attendance at programme sessions and baseline characteristics of caregivers and their families. The second stage consisted of estimating programme effects while also accounting for variation in attendance, achieved by comparing CACE to ITT effects.

Predicting Attendance

Thailand. Results of the univariate logistic regression analyses, which included each predictor variable one at a time, are presented as unadjusted coefficients and IRRs in Table 6. Only two predictors were strongly associated to caregiver attendance in the univariate analyses, namely caregiver education and caregiver relationship to the target child. Specifically, caregivers with higher education attended 8% less than caregivers with lower education levels (IRR = 0.92 [95% CI = 0.88, 0.97], $p = 0.003$). Attendance was 25% lower for mothers compared to grandparents and great-grandparents (IRR = 0.75, 95% CI [0.61, 0.92], $p = 0.006$).[3]

There were also associations between attendance and government benefits, caregiver depression, proactive parenting, and caregiver age. However, all these associations were predominantly weak and should be treated with caution since their confidence intervals are approaching 1.00. Attendance was 32% higher for caregivers who received government benefits (IRR = 1.32, 95% CI [1.00, 1.79], $p = 0.063$), although the confidence intervals suggest an imprecise association. A precise yet weak effect suggests that caregivers with

[3] Only three great-grandparents were allocated to the intervention arm. Thus, a categorical variable with two categories was created (0 = grandparent/great-grandparent; 1 = parent).

higher levels of depression attended 2% more sessions (IRR = 1.02, 95% CI [1.00, 1.03], p = 0.077). Similarly, again a precise yet weak effect suggests that caregivers with higher levels of proactive parenting attended 1% more sessions (IRR = 1.01, 95% CI [1.00, 1.03], p = 0.119). Although grandparents and great-grandparents were significantly more likely to attend (see above), it appears as though older caregivers only attended 1% more sessions (IRR = 1.01, 95% CI [1.00, 1.02], p = 0.007) which is near to no effect.

An inspection of IRRs and their 95% CIs indicate that there were no significant associations between attendance and the following variables: household assets; household income; household hunger; household size; caregiver health; caregiver history of abuse; overall caregiver mental health problems well as caregiver anxiety and stress; intimate partner violence; caregiver disability; overall child maltreatment as well as physical and emotional abuse subscales; overall positive parenting and the positive parenting and limit setting subscales; Parent Daily Report for child and parenting behaviour; ECBI child problem behaviour and child problem behaviour intensity; child age and gender; marital status; and caregiver unemployment.

In line with Hosmer and Lemeshow's (2000) recommendations on retaining predictors with associations smaller or equal to .25, the following variables were retained for multivariable modelling: caregiver education (p = 0.003), government benefits (p = 0.063), household assets (p = 0.161), caregiver depression (p = 0.077), caregiver disability (p = 0.249), proactive parenting (p = 0.119), parent daily report parenting behaviour (p = 0.166), caregiver age (p = 0.007), and caregiver relationship to child (p = 0.006).

Table 6

Univariate Logistic Regression Analyses Predicting Attendance: Thailand

Predictors	Estimate	S.E.	IRR [95% CI]	*p*
Economic and educational				
Caregiver education	**-0.08**	**0.03**	**0.92 [0.88, 0.97]**	**0.003**
Government benefits (yes)	0.28	0.14	1.32 [1.00, 1.79]	0.063
Low income	0.25	0.31	1.29 [0.68, 2.31]	0.415
Middle income	0.16	0.19	1.17 [0.81, 1.73]	0.409
Upper middle income	0.14	0.18	1.15 [0.82, 1.67]	0.448
Household assets	-0.02	0.02	0.98 [0.94, 1.01]	0.161
Household hunger	0.02	0.03	1.02 [0.97, 1.07]	0.413
Household size	-0.01	0.03	0.99 [0.93, 1.06]	0.829
Social and Health				
Caregiver health	0.00	0.01	1.01 [0.91, 1.12]	0.907
ICAST-R (history of abuse)	-0.01	0.02	0.99 [0.95, 1.02]	0.474
DASS Total	0.01	0.01	1.01 [0.99, 1.02]	0.307
DASS Depression	0.02	0.01	1.02 [1.00, 1.03]	0.077
DASS Anxiety	0.00	0.01	1.00 [0.98, 1.02]	0.929
DASS Stress	0.01	0.01	1.01 [0.99, 1.02]	0.425
CTS2S Intimate partner violence	0.03	0.04	1.03 [0.95, 1.10]	0.470
Caregiver has a disability	0.12	0.11	1.13 [0.91, 1.39]	0.249
Parenting and child behaviour				
ICAST Total	-0.00	0.00	1.00 [0.99, 1.01]	0.522
ICAST Physical abuse	-0.01	0.01	0.99 [0.97, 1.01]	0.497
ICAST Emotional abuse	-0.00	0.01	1.00 [0.98, 1.01]	0.631
PARYC Total	0.00	0.00	1.00 [1.00, 1.01]	0.377
PARYC Positive parenting	0.00	0.01	1.00 [0.98, 1.02]	0.861
PARCY Setting limits	0.00	0.01	1.00 [0.99, 1.02]	0.853
PARYC Proactive parenting	0.01	0.01	1.01 [1.00, 1.03]	0.119
PS Over-reactivity	0.04	0.05	1.04 [0.94, 1.15]	0.420
APQ Poor monitoring/supervision	-0.01	0.01	0.99 [0.97, 1.02]	0.621
PDR Parent behaviour	-0.06	0.04	0.94 [0.86, 1.03]	0.166
PDR Child behaviour	0.01	0.01	1.01 [0.99, 1.02]	0.509
ECBI Problem	0.00	0.01	1.00 [0.99, 1.02]	0.700
ECBI Intensity	0.00	0.00	1.00 [1.00, 1.00]	0.523
Sociodemographic characteristics				
Caregiver age	0.01	0.00	1.01 [1.00, 1.02]	0.007
Child age	0.01	0.03	1.01 [0.96, 1.06]	0.817
Child gender, boy	-0.06	0.10	0.95 [0.77, 1.16]	0.590
Marital status, separated	0.22	0.22	1.25 [0.80, 1.92]	0.318
Marital status, unmarried	0.10	0.14	1.11 [0.84, 1.47]	0.479
Biological parent	**-0.28**	**0.10**	**0.75 [0.61, 0.92]**	**0.006**
Caregiver unemployed	-0.021	0.10	0.98 [0.80, 1.20]	0.837

Note. Only self-report measures were included in the univariable analyses. SE = unadjusted standard error;

IRR = Incidence rate ratio (IRR=1.00 no effect, IRR < 1.00 lower attendance, IRR > 1.00 higher attendance);

95 Cl = 95% confidence intervals. ICAST-R = International Society for the Prevention of Child Abuse and

Neglect Child Abuse Screening Tool Retrospective version; ICAST = International Society for the Prevention of Child Abuse and Neglect Child Abuse Screening Tool; PARYC = Parenting Young Children Scale; PS = Parenting Scale; APQ = Alabama Parenting Questionnaire; DASS = Depression, Anxiety, and Stress Scale; ECBI = Eyberg Child Behaviour Inventory; PDR = Parent Daily Report Checklist; CTS-2S = Revised Conflict Tactics Scale Short Form.

An inspection of Spearman correlation coefficients (see Table 7) of all the variables selected for multivariable modelling revealed that there was a strong positive correlation between caregiver age and caregiver relationship to the child ($r = .74$, $p < 0.01$). However, the variance inflation factor indicated no multicollinearity within the multivariable model, with all values below 5. Caregiver age and child relationship were therefore both retained in the multivariate model. To avoid inclusion of two socioeconomic status variables in a single model, receiving government benefits and number of household assets were entered into two separate models. To identify the most parsimonious fit, a partial likelihood ratio test indicated that the model with household assets was significantly better at predicting attendance ($X^2 = 0.25$, $df = 12$, $p < 0.001$) and, consequently, the government benefits variable was not retained as a predictor.

Table 7

Correlation Matrix: Variables Identified for Multivariate Analysis, Thailand

	1	2	3	4	5	6	7	8	9	10
1. Caregiver age	_____									
2. Relationship (grandparent)	0.74**	_____								
3. PDR Parent behaviour[1]	-0.05	-0.18	_____							
4. PARYC Proactive parenting[2]	0.15	0.18	-0.12	_____						
5. Caregiver has a disability	0.34**	0.27*	-0.20	0.23	_____					
6. DASS Depression[3]	0.34**	0.30*	-0.19	0.11	0.48**	_____				
7. Total household assets	-0.00	-0.13	0.08	-0.03	0.07	-0.04	_____			
8. Government benefits	0.19	0.03	-0.10	0.07	0.10	0.22	-0.36*	_____		
9. Caregiver education	-0.58**	-0.56**	0.15	-0.49	-0.30*	-0.26*	0.28*	-0.26*	_____	
10. Overall attendance	0.42**	0.32*	-0.21	0.17	0.04	0.21	-0.19	0.14	-0.43**	_____

Note. Spearman's Correlation Coefficients. *Correlation is significant at the 0.05 level (2-tailed); **Correlation is significant at the 0.01 level (2-tailed).

[1]Parent Daily Report Checklist; [2]Parenting Young Children Scale; [3]Depression, Anxiety, and Stress Scale.

The results of the final multivariable model are presented in Table 8**.** They indicate that education was a strong predictor of the number of sessions attended. Caregivers with higher levels of education were estimated to attend fewer sessions - specifically, caregivers with higher educational backgrounds attended 7% fewer sessions than those with lower education (IRR = 0.93, 95% CI [0.87, 1.00], $p = 0.011$).

As already identified in the univariable models, there appear to be weak associations between caregiver age and proactive parenting and attendance. Older caregivers had a slightly higher chance of attending, with one additional year of age predicting 1% more sessions attended (IRR = 1.01, 95% CI [1.00, 1.02], $p = 0.022$). Caregivers with higher proactive parenting also attended 1% more sessions (IRR = 1.01, 95% CI [1.00, 1.03], $p = 0.061$). However, these results are somewhat meaningless given that the small effect sizes and the 95% CI are approaching 1.00.

Results also indicate that biological mothers attended 11% fewer sessions than non-mothers; however, this association was no longer significant in the multivariable model as indicted by the overlap of the 95% CIs (IRR = 0.89, 95% CI [0.78, 1.01], $p = 0.157$). There was no significant association between attendance and household assets, caregiver depression, caregiver disability, and parent daily report parenting behaviour.

The appropriateness of the final multivariable model was confirmed through goodness of fit tests and model diagnostics. Neither the Hosmer and Lemeshow goodness of fit test ($X^2 = -1.51$, $df = 8$, $p = 1.00$) or the Pearson's Chi-squared residuals ($p = 0.989$) indicated a lack a of model fit. The relationship between the observed counts and the fitted values was reasonably good (see Appendix I), and there were no clear patterns in the qq-plot or the residuals that would indicate non-linearity or overdispersion. To further investigate this, the estimated dispersion parameter, calculated using the model residual degrees of freedom as a measure of the mean and the deviance as a measure variance, confirmed the Poisson

regression assumption that the variance equals the mean (dispersion parameter = 1.02). Finally, continuous predictors did not display any major deviations from linearity (see Appendix J). The variance inflation factors were no greater than 2.31 which is well below the recommended threshold of 5.

Table 8

Final Multivariable Model of Predicting Attendance: Thailand

Predictors	Estimate	Robust S.E.	*p*	IRR	IRR 95 CI
(Intercept)	1.89	0.38	<0.001	6.64	[2.32, 18.67]
Economic and educational					
Caregiver education	**-0.07**	**0.03**	**0.011**	**0.93**	**[0.87, 1.00]**
Household assets	-0.01	0.01	0.299	0.99	[0.95, 1.02]
Social and Health					
DASS Depression	0.01	0.01	0.397	1.01	[0.98, 1.03]
Caregiver has a disability	-0.12	0.07	0.073	0.89	[0.68, 1.16]
Parenting and child behaviour					
PARYC Proactive parenting	**0.01**	**0.01**	**0.061**	**1.01**	**[1.00, 1.03]**
PDR Parent behaviour	-0.03	0.04	0.370	0.97	[0.88, 1.07]
Sociodemographic characteristics					
Caregiver age	**0.01**	**0.01**	**0.022**	**1.01**	**[1.00, 1.02]**
Biological parent	-0.12	0.08	0.157	0.89	[0.78, 1.01]

Note. Estimates for parenting group are not shown but were included in the analysis to account for dependence in the data. Robust sandwich estimators were used to estimate coefficients, standard errors (S.E.), and *p*-values. IRR = Incidence rate ratio (IRR=1.00 no effect, IRR < 1.00 lower attendance, IRR > 1.00 higher attendance); 95 Cl = 95% confidence intervals. DASS = Depression, Anxiety, and Stress Scale; PARYC = Parenting Young Children Scale; PDR = Parent Daily Report Checklist.

Philippines. Results of the univariate logistic regression analyses are presented as unadjusted coefficients and IRRs in Table 9. Three predictors were strongly associated with caregiver attendance, including child gender, caregiver health, and intimate partner violence. Specifically, caregivers with boys attended 31% more sessions than caregivers with girls (IRR = 1.31, 96% CI [1.09, 1.58], $p = 0.005$). Attendance among healthier caregivers was 13% lower than among less healthy caregivers (IRR = 0.87, 95% CI [0.76, 1.00], $p = 0.048$). Caregivers who experienced higher rates of intimate partner violence attended 7% fewer sessions (IRR = 0.93, 95% CI [0.89, 0.97], $p = 0.002$).

Much weaker effects were found for child maltreatment, child emotional abuse, and caregiver over-reactivity. However, these results should be interrupted with caution since the size of effects are small, and the IRRs and CIs were close to 1.00. A weak yet precise effect suggests that caregivers who reported perpetrating higher rates of child maltreatment attended 1% more sessions (IRR = 1.01, 95% CI [1.00, 1.01], $p = 0.159$). Further, a weak but precise effect suggests that higher emotional abuse by caregivers was associated with 2% more attendance (IRR = 1.02, 95% CI [1.00, 1.03], $p = 0.057$). Caregiver over-reactivity was also associated with a weak effect on attendance, with caregivers reporting higher over-reactivity attending 1% more sessions (IRR = 1.01, 95% CI [1.00, 1.02], $p = 0.159$).

With IRR 95% CIs overlapping 1.00, no significant associations were found between attendance and caregiver education level; household assets; household hunger level; household size; caregiver history of abuse; caregiver depression, anxiety, and stress; caregiver disability; child physical abuse; positive parenting; dysfunctional parenting and the laxness subscale; PDR on parenting and child behaviour; ECBI child problem behaviour and child problem behaviour intensity; caregiver and child age; marital status; and caregiver employment status.

Following Hosmer and Lemeshow's (2000) guidelines for variable selection, associations with *p*-values smaller than or equal to .25 were retained for multivariable modelling. These included: household assets ($p = 0.172$), household size ($p = 0.047$), adult health ($p = 0.048$), intimate partner violence ($p = 0.002$), overall child maltreatment ($p = 0.159$), child emotional abuse ($p = 0.057$), overall positive parenting ($p = 0.184$) as well as the positive parenting subscale ($p = 0.017$), caregiver over-reactivity ($p = 0.159$), parent daily report parenting and child behaviour ($p = 0.240$; $p = 0.119$, respectively), and child gender ($p = 0.005$).

An inspection of Spearman correlation coefficients (see Table 10) of all the variables selected for the multivariable model revealed that there was a strong positive correlation between overall positive parenting (PARYC total) and the Positive Parenting subscale ($r = .74$, $p < 0.01$). Since the other PARYC subscales (Proactive Parenting and Setting Limits) were not strong predictors of attendance, the Positive Parenting total scale was not retained. Similarly, there was also a strong positive correlation between overall child maltreatment (ICAST total) and the Emotional Abuse subscale ($r = 0.90$, $p < 0.01$), and so only the emotional abuse subscale was retained in the multivariable model. Next, household size and household assets were entered into two separate models to identify which of these socioeconomic measures was the most parsimonious fit. A partial likelihood ratio test indicated that the model containing household size was significantly better ($X^2 = 1.35$, $df = 14$, $p < 0.001$). The household assets variable was therefore not retained.

Table 9

Univariate Logistic Regression Analyses Predicting Attendance: Philippines

Predictors	Estimate	S.E.	IRR [95% CI]	*p*
Economic and educational				
Caregiver education	0.02	0.03	1.02 [0.97, 1.07]	0.500
Household assets	-0.05	0.03	0.95 [0.89, 1.02]	0.172
Household hunger	-0.01	0.02	0.99 [0.94, 1.03]	0.590
Household size	-0.04	0.02	0.96 [0.92, 1.00]	0.047
Social and Health				
Caregiver health	**-0.14**	**0.07**	**0.87 [0.76, 1.00]**	**0.048**
ICAST-R history of abuse	-0.01	0.05	0.99 [0.90, 1.09]	0.815
DASS Total	0.00	0.01	1.00 [0.99, 1.02]	0.591
DASS Depression	0.00	0.01	1.00 [0.98, 1.03]	0.715
DASS Anxiety	0.00	0.01	1.00 [0.98, 1.02]	0.733
DASS Stress	0.01	0.01	1.01 [0.99, 1.02]	0.398
CTS2S intimate partner violence	**-0.07**	**0.02**	**0.93 [0.89, 0.97]**	**0.002**
Caregiver has a disability	0.08	0.10	1.09 [0.90, 1.32]	0.385
Parenting and child behaviour				
ICAST Total	0.01	0.00	1.01 [1.00, 1.01]	0.159
ICAST Physical abuse	0.01	0.01	1.01 [0.99, 1.03]	0.253
ICAST Emotional abuse	0.02	0.01	1.02 [1.00, 1.03]	0.057
PARYC Total	-0.01	0.00	0.99 [0.99, 1.00]	0.184
PARYC Positive parenting	-0.02	0.01	0.98 [0.96, 1.00]	0.017
PARYC Setting limits	-0.00	0.01	1.00 [0.98, 1.02]	0.798
PARYC Proactive parenting	-0.00	0.01	1.00 [0.98, 1.01]	0.713
PS Total (dysfunctional parenting)	0.00	0.00	1.00 [1.00, 1.01]	0.417
PS Laxness	-0.00	0.01	1.00 [0.99, 1.01]	0.804
PS Over-reactivity	0.01	0.01	1.01 [1.00, 1.02]	0.159
PDR Parent behaviour	0.05	0.04	1.05 [0.97, 1.13]	0.240
PDR Child behaviour	-0.02	0.01	0.98 [0.96, 1.00]	0.119
ECBI Problem	-0.01	0.01	0.99 [0.98, 1.01]	0.409
ECBI Intensity	-0.00	0.00	1.00 [0.99, 1.00]	0.560
Sociodemographic characteristics				
Caregiver age	0.01	0.01	1.01 [0.99, 1.02]	0.287
Child age	-0.04	0.04	0.96 [0.90, 1.04]	0.312
Child gender, boy	**0.27**	**0.10**	**1.31 [1.09, 1.58]**	**0.005**
Marital status, separated	-0.26	0.24	0.77 [0.49, 1.27]	0.271
Marital status, unmarried	-0.25	0.24	0.78 [0.50, 1.20]	0.295
Caregiver unemployed	0.07	0.10	1.08 [0.88, 1.32]	0.481

Note. Caregiver gender and child relationship variables were not included as only females attended and almost all caregivers were biological mothers. Estimates = unadjusted coefficients; SE = unadjusted standard error; IRR = Incidence rate ratio (IRR=1.00 no effect, IRR < 1.00 lower attendance, IRR > 1.00 higher attendance); 95 Cl = 95% confidence intervals. ICAST-R =

International Society for the Prevention of Child Abuse and Neglect Child Abuse Screening Tool Retrospective version; ICAST = International Society for the Prevention of Child Abuse and Neglect Child Abuse Screening Tool; PARYC = Parenting Young Children Scale; PS = Parenting Scale; APQ = Alabama Parenting Questionnaire; DASS = Depression, Anxiety, and Stress Scale; ECBI = Eyberg Child Behaviour Inventory; PDR = Parent Daily Report Checklist; CTS-2S = Revised Conflict Tactics Scale Short Form.

Table 10

Correlation Matrix: Variables Identified for Multivariate Analysis, Philippines

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Child gender	_____												
2. PDR child behaviour[1]	-0.03	_____											
3. PDR parent behaviour[1]	0.05	-0.47**	_____										
4. PS overreactive parenting[2]	-0.12	0.06	-0.19	_____									
5. PARYC total[3]	-0.05	-0.04	0.06	0.02	_____								
6. PARYC positive parenting[3]	-0.12	0.08	-0.06	0.04	0.74**	_____							
7. ICAST total[4]	-0.12	-0.00	0.04	0.29*	0.12	0.06	_____						
8. ICAST emotional abuse[4]	-0.03	0.09	0.00	0.21	0.12	0.11	0.90**	_____					
9. CTS2S[5]	-0.15	0.39**	-0.42**	0.03	0.01	0.09	0.35**	0.35**	_____				
10. Caregiver health	-0.20	-0.02	0.32*	-0.02	-0.02	-0.02	-0.11	-0.08	-0.28*	_____			
11. Household size	0.08	-0.24	0.23	-0.16	-0.17	-0.27*	0.04	-0.02	-0.14	-0.06	_____		
12. Household assets	-0.11	-0.15	0.12	0.10	0.18	0.22	-0.08	-0.01	-0.18	0.15	0.02	_____	
13. Overall attendance	0.34**	-0.08	0.12	0.10	-0.15	-0.21	0.18	0.19	-0.23	-0.14	-0.13	-0.08	_____

Note. Spearman's Correlation Coefficients. *Correlation is significant at the 0.05 level (2-tailed); **Correlation is significant at the 0.01 level (2-tailed). [1]Parent Daily Report Checklist; [2]Parenting Scale; [3]Parenting Young Children Scale; [4]International Society for the Prevention of Child Abuse and Neglect Child Abuse Screening Tool; [5]Revised Conflict Tactics Scale Short Form.

The results of the final multivariable model are presented in Table 11**.** As suggested in the univariate analyses, child gender, caregiver health, intimate partner violence, and child emotional abuse also showed unique relationships with attendance in the multivariable model. Specifically, attendance was 25% higher among caregivers with boys compared to caregivers with girls (IRR = 1.25, [1.00, 1.56], $p = 0.032$). Attendance was 15% lower among caregivers who reported higher levels of health (IRR = 0.85, 95% CI [0.72, 0.99], $p = 0.015$). Caregivers who experienced more intimate partner violence had a slightly lower attendance, with a one-unit increase in intimate partner violence exposure associated with 8% fewer sessions attended (IRR = 0.92, 95% CI [0.87, 0.98], $p = 0.021$). Caregivers who reported higher levels of child emotional abuse attended 3% more sessions (IRR = 1.03, 95% CI [1.01, 1.04], $p < 0.001$). Based on the IRRs and 95% CI, there was little to no indication that household size, positive parenting, parent daily report for child and parenting behaviour, and over-reactive parenting were associated with attendance in the multivariable model.

Model fit was confirmed through a combination of goodness of fit tests and model diagnostics. Neither the Hosmer and Lemeshow goodness of fit test ($X^2 = -1.68$, $df = 8$, $p = 1.00$) or the Pearson's Chi-squared residuals test ($p = 0.134$) provided evidence of lack of model fit. The relationship between the fitted values and the observed counts did not identify model misfit (see Appendix K). There were no obvious patterns in the qq-plot or the residuals that would suggest overdispersion or non-linearity, and continuous predictors did not display any major deviations from linearity (see Appendix L). The variance inflation factors, measuring multicollinearity in the model, were no greater than 1.85 which is well below the recommended threshold of 5.00.

Table 11

Final Multivariable Model Predicting Attendance: Philippines

Predictors	Estimate	Robust S.E.	*p*	IRR	IRR 95 CI
(Intercept)	3.04	0.77	<0.001	20.89	[5.17, 83.94]
Economic and educational					
Household size	-0.05	0.03	0.097	0.95	[0.91, 1.00]
Social and Health					
Caregiver health	**-0.17**	**0.07**	**0.015**	**0.85**	**[0.72, 0.99]**
CTS2S intimate partner violence	**-0.08**	**0.03**	**0.021**	**0.92**	**[0.87, 0.98]**
Parenting and child behaviour					
ICAST Emotional abuse	**0.03**	**0.01**	**<0.001**	**1.03**	**[1.01, 1.04]**
PARYC Positive parenting	-0.02	0.01	0.048	0.98	[0.96, 1.00]
PS Over-reactivity	0.00	0.01	0.483	1.00	[0.99, 1.02]
PDR Parent behaviour	0.04	0.05	0.352	1.04	[0.95, 1.15]
PDR Child behaviour	-0.01	0.01	0.700	0.99	[0.97, 1.02]
Sociodemographic characteristics					
Child gender, boy	**0.22**	**0.10**	**0.032**	**1.25**	**[1.00, 1.56]**

Note. Estimates for parenting group are not shown but were included in the analysis to account for dependence in the data. Robust sandwich estimators were used to estimate coefficients, standard errors (S.E.), and p-values. IRR = Incidence rate ratio (IRR=1.00 no effect, IRR < 1.00 lower attendance, IRR > 1.00 higher attendance); 95 Cl = 95% confidence intervals. ICAST = International Society for the Prevention of Child Abuse and Neglect Child Abuse Screening Tool; PARYC = Parenting Young Children Scale; PS = Parenting Scale; PDR = Parent Daily Report Checklist; CTS-2S = Revised Conflict Tactics Scale Short Form.

Impact of Attendance on Primary and Secondary Outcomes

The second stage of the data analysis involved determining the contribution of caregivers' attendance to programme outcomes. Analyses in this section predominantly focus on the current study's primary outcomes (i.e., child maltreatment), but also briefly discusses results for the secondary outcomes.

First, overall programme effects were estimated via ITT analyses. Second, CACE were estimated for the two definitions of compliance (i.e., moderately attendance = attending at least 50% of sessions; high attendance = attending at least 75% of sessions) to account for varying levels of attendance in the programmes. CACE estimates were then compared to ITT estimates and sensitivity analyses, which included covariates, were conducted to examine the robustness of CACE results. As discussed in the data analysis section above, logistic regression analyses were conducted for non-normally distributed data and linear regression analyses were conducted for data approximating normality. Hence, IRRs are only reported for logistic regression models whereas unstandardised coefficients are reported for linear models.

Prior to model fitting, descriptive statistics were inspected for each measure - both by allocation as well as for the two definitions of compliance. Tables 12 to 15 summarise these descriptive statistics for primary and secondary outcome measures for both RCTs. Overall, a general trend was visible across both RCTs, where control group means for child maltreatment, physical abuse, and emotional abuse at post-test and follow-up were substantially higher than those for the intervention group. In addition, participants who attended at least 75% of the programmes had the lowest means, suggesting that higher attendance may indeed influence outcomes. Similar tends were also evident for secondary outcomes, with participants who attended more sessions having higher means for positive parenting and lower means for over-reactive parenting, child behaviour problems, and caregiver mental health issues after the programme.

Table 12

Descriptive Statistics of Primary Outcomes: Thailand

	Intervention		Control		Attend ≥ 4 sessions		Attend ≥ 6 sessions	
Outcome	*M* (*SD*)	*Mdn*	*M* (*SD*)	*Mdn*	*M* (*SD*)	*Mdn*	*M* (*SD*)	*Mdn*
HOME Harsh parenting								
Baseline	1.17 (1.04)	1.00	1.26 (1.11)	1.00	1.15 (1.04)	1.00	1.13 (1.03)	1.00
Posttest	0.55 (0.65)	0.00	1.02 (1.07)	1.00	0.62 (0.69)	1.00	0.64 (0.70)	1.00
Follow-up	0.56 (0.71)	0.00	1.07 (0.92)	1.00	0.60 (0.72)	0.00	0.57 (0.69)	0.00
ICAST Overall								
Baseline	9.78 (10.84)	6.00	11.48 (13.65)	8.00	9.20 (9.76)	6.00	9.43 (10.23)	6.00
Posttest	3.17 (4.67)	1.00	9.00 (10.81)	5.00	3.02 (4.34)	1.00	2.72 (4.23)	1.00
Follow-up	2.79 (3.91)	1.00	7.21 (9.95)	4.00	2.62 (3.86)	0.50	2.63 (4.05)	0.00
ICAST Physical abuse								
Baseline	4.24 (5.53)	2.00	5.70 (7.80)	3.00	4.11 (5.26)	2.00	4.13 (5.43)	2.00
Posttest	1.38 (2.54)	0.00	4.16 (5.90)	2.00	1.19 (2.03)	0.00	0.98 (1.79)	0.00
Follow-up	0.98 (2.27)	0.00	2.96 (4.85)	1.00	0.96 (2.20)	0.00	0.93 (2.28)	0.00
ICAST Emotional abuse								
Baseline	5.54 (6.55)	3.00	5.77 (7.80)	4.00	5.09 (5.67)	3.00	5.30 (5.98)	3.00
Posttest	1.79 (2.77)	0.00	4.84 (6.38)	3.00	1.83 (2.83)	0.00	1.74 (2.86)	0.00
Follow-up	1.81 (2.71)	0.00	4.25 (5.99)	2.00	1.65 (2.56)	0.00	1.70 (2.66)	0.00

Note. M = mean; SD = standard deviation; Mdn = median. HOME = The Home Observation for Measurement of Environment Inventory Abusive Harsh Parenting subscale; ICAST = International Society for the Prevention of Child Abuse and Neglect Child Abuse Screening Tool.

Table 13

Descriptive Statistics of Secondary Outcomes: Thailand

	Intervention		Control		Attend ≥ 4 sessions		Attend ≥ 6 sessions	
Outcome	*M* (*SD*)	*Mdn*	*M* (*SD*)	*Mdn*	*M* (*SD*)	*Mdn*	*M* (*SD*)	*Mdn*
PARYC Total								
Baseline	68.31 (15.00)	69.00	68.33 (14.86)	69.00	68.85 (15.39)	69.00	69.19 (15.24)	69.00
Posttest	77.91 (13.65)	79.50	70.41 (17.92)	75.50	78.15 (13.83)	80.00	79.53 (13.65)	80.00
Follow-up	77.53 (13.37)	78.00	70.72 (16.11)	72.00	77.69 (13.78)	78.00	78.15 (13.60)	78.00
APQ Poor monitoring								
Baseline	15.63 (3.74)	15.00	16.56 (5.46)	15.00	15.56 (3.81)	15.00	15.79 (3.94)	15.00
Posttest	13.55 (2.94)	12.00	16.03 (5.72)	14.00	13.51 (2.87)	12.00	13.77 (2.94)	13.00
Follow-up	13.74 (3.65)	13.00	15.81 (5.70)	14.00	13.81 (3.69)	13.00	13.96 (3.90)	13/00
PS Over-reactivity								
Baseline	2.41 (0.96)	2.20	2.76 (1.14)	2.60	2.44 (0.98)	2.30	2.44 (1.04)	2.30
Posttest	1.78 (0.74)	1.65	2.31 (1.08)	2.20	1.79 (0.76)	1.70	1.83 (0.79)	1.70
Follow-up	1.81 (0.75)	1.80	2.24 (1.04)	2.30	1.78 (0.75)	1.70	1.80 (0.76)	1.70
ECBI Problem								
Baseline	5.07 (7.49)	1.00	4.85 (6.71)	1.00	5.04 (7.10)	2.00	4.98 (7.20)	2.00
Posttest	2.36 (5.42)	0.00	4.02 (7.24)	0.00	2.12 (5.07)	0.00	1.94 (5.02)	0.00
Follow-up	1.72 (3.96)	0.00	3.67 (7.19)	0.00	1.65 (3.92)	0.00	1.67 (4.03)	0.00
ECBI Intensity								
Baseline	102.56 (31.92)	96.00	104.97 (35.92)	101.00	104.06 (31.14)	96.00	103.45 (32.33)	96.00
Posttest	82.14 (30.51)	75.00	101.48 (37.60)	94.00	81.92 (30.96)	75.00	81.32 (31.83)	71.00
Follow-up	81.74 (28.54)	76.00	96.23 (38.76)	97.00	82.44 (28.92)	76.50	83.37 (30.01)	78.00
DASS Total								
Baseline	9.15 (7.48)	7.00	8.54 (8.59)	6.00	9.43 (7.45)	7.50	9.64 (7.63)	8.00
Posttest	5.28 (5.96)	3.00	8.12 (7.24)	6.00	5.40 (6.14)	3.00	5.49 (6.38)	3.00
Follow-up	4.65 (6.03)	3.00	7.26 (7.68)	6.00	4.81 (6.16)	3.00	5.02 (6.47)	2.50

Note. M = mean; SD = standard deviation; Mdn = median. PARYC = Parenting Young Children Scale; PS = Parenting Scale; APQ = Alabama Parenting Questionnaire; ECBI = Eyberg Child Behaviour Inventory; DASS = Depression, Anxiety, and Stress Scale.

Table 14

Descriptive Statistics of Primary Outcomes: Philippines

	Intervention		Control		Attend ≥ 6 sessions		Attend ≥ 9 sessions	
Outcome	*M* (*SD*)	*Mdn*	*M* (*SD*)	*Mdn*	*M* (*SD*)	*Mdn*	*M* (*SD*)	*Mdn*
ICAST Overall								
Baseline	12.45 (12.00)	9.00	14.07 (15.45)	9.50	13.49 (13.08)	10.00	14.58 (13.23)	11.00
Post-test	6.12 (5.27)	4.00	13.79 (16.90)	8.50	6.28 (5.52)	4.00	7.92 (5.98)	7.50
Follow-up	7.32 (7.54)	5.00	11.54 (13.17)	7.00	6.45 (5.99)	4.50	6.68 (5.12)	5.00
ICAST Physical abuse								
Baseline	3.37 (4.38)	2.00	4.03 (5.29)	3.00	3.79 (4.85)	3.00	3.92 (5.24)	3.50
Post-test	1.36 (2.07)	1.00	3.64 (5.49)	1.50	1.49 (2.26)	1.00	1.85 (2.49)	1.00
Follow-up	1.98 (3.16)	1.00	3.30 (4.57)	2.00	1.55 (2.41)	0.50	1.64 (2.33)	1.00
ICAST Emotional abuse								
Baseline	6.35 (5.24)	5.00	7.18 (7.29)	5.00	6.81 (5.70)	5.00	7.77 (5.89)	6.50
Post-test	3.53 (3.35)	2.00	7.36 (8.39)	5.00	3.58 (3.44)	3.00	4.35 (4.05)	3.50
Follow-up	3.95 (4.20)	3.00	5.88 (6.31)	4.00	3.50 (3.53)	2.00	3.44 (2.97)	3.00

Note. M = mean; SD = standard deviation; Mdn = median. ICAST = International Society for the Prevention of Child Abuse and Neglect Child Abuse Screening Tool.

Table 15

Descriptive Statistics of Secondary Outcomes: Philippines

	Intervention		Control		Attend ≥ 6 sessions		Attend ≥ 9 sessions	
Outcome	*M* (*SD*)	*Mdn*	*M* (*SD*)	*Mdn*	*M* (*SD*)	*Mdn*	*M* (*SD*)	*Mdn*
PARYC Total								
Baseline	102.28 (11.04)	101.50	101.70 (13.20)	100.50	101.63 (11.83)	99.00	100.65 (9.50)	99.50
Posttest	103.50 (14.22)	103.00	99.28 (11.13)	97.00	103.58 (13.73)	102.00	105.65 (15.64)	103.00
Follow-up	105.93 (13.19)	103.00	104.00 (13.35)	101.00	106.43 (12.81)	103.00	110.48 (14.59)	107.00
PS Dysfunctional parenting								
Baseline	112.10 (13.41)	113.50	108.93 (15.05)	110.50	113.49 (12.37)	114.00	110.73 (12.73)	108.50
Posttest	102.02 (11.85)	103.00	111.02 (10.59)	110.50	101.49 (12.31)	102.00	97.96 (10.62)	98.00
Follow-up	105.14 (14.54)	104.50	107.49 (15.41)	106.50	101.74 (14.43)	100.00	97.84 (14.52)	96.00
ECBI Problem								
Baseline	6.80 (6.43)	4.50	7.68 (7.99)	5.00	6.58 (6.25)	4.00	6.27 (6.50)	3.50
Posttest	6.72 (6.95)	4.50	8.71 (9.54)	6.50	6.00 (6.83)	3.00	5.31 (7.44)	1.00
Follow-up	7.23 (7.86)	5.50	7.56 (9.36)	3.00	7.24 (8.22)	5.50	7.56 (8.45)	7.00
ECBI Intensity								
Baseline	120.00 (21.59)	118.00	123.93 (21.82)	127.50	118.84 (20.94)	115.00	119.15 (20.60)	119.00
Posttest	114.36 (23.19)	118.00	124.34 (24.95)	130.00	112.49 (23.96)	116.00	113.31 (27.84)	120.00
Follow-up	114.27 (23.43)	119.50	120.32 (23.74)	125.00	115.05 (23.00)	118.00	117.36 (22.37)	120.00
DASS Total								
Baseline	9.25 (6.02)	8.00	13.55 (8.37)	13.00	9.70 (6.10)	9.00	9.38 (5.51)	8.50
Posttest	10.41 (6.96)	9.50	13.41 (8.65)	11.50	9.98 (6.64)	9.00	9.65 (6.60)	7.50
Follow-up	10.77 (7.61)	11.00	12.54 (7.40)	7.40	10.49 (8.03)	11.00	8.62 (5.82)	6.00

Note. M = mean; SD = standard deviation; Mdn = median. PARYC = Parenting Young Children Scale; PS = Parenting Scale; ECBI = Eyberg Child Behaviour Inventory; DASS = Depression, Anxiety, and Stress Scale.

Intention-To-Treat Analyses: Thailand. Programme effects, where the control group was specified as the reference variable, are reported in the left-hand columns of Table 16 (primary outcomes) and Table 17 (secondary outcomes). Controlling for baseline scores of each outcome, ITT estimates suggest that caregivers who received the intervention had 44% reduced risk of harsh parenting practices, as measured by the HOME harsh parenting observational assessment, at post-test (IRR = 0.56, 95% CI [0.36, 0.86], $p = 0.003$) and 46% reduced risk at follow-up (IRR = 0.54, 95% CI [0.34, 0.82], $p = 0.001$). Analyses also showed that caregivers in the intervention group reported 63% reduced risk of overall child maltreatment at post-test (IRR = 0.37, 95% CI [0.32, 0.45], $p < 0.001$) and 59% reduced risk at follow-up (IRR = 0.41, 95% CI [0.34, 0.50], $p < 0.001$). Significant reductions in the Physical and Emotional Abuse subscales were also evident. For physical abuse, caregivers reported 62% reduced risk at post-test (IRR = 0.38, 95% CI [0.30, 0.49], $p < 0.001$) and 60% reduced risk at follow-up (IRR = 0.40, 95% CI [0.29, 0.53], $p = 0.013$). Caregivers also reported 62% reduced risk of emotional abuse at post-test (IRR = 0.38, 95% CI [0.30, 0.47], $p < 0.001$) and 56% reduced risk at follow-up (IRR = 0.40, 95% CI [0.34, 0.55], $p = 0.002$).

Among the secondary outcomes, treatment allocation also appeared to be a significant predictor of increased positive parenting, and decreased over-reactive parenting, child problem behaviour, and caregiver mental health issues. Regarding positive parenting, caregivers in the intervention group reported significant increases in overall positive parenting based on the PARYC total scale at post-test (unstandardised coefficient [b] = 7.46, 95% CI [2.92, 11.99], $p = 0.001$) and at follow-up ($b = 6.61$, 95% CI [2.31, 10.92], $p = 0.003$). Caregivers also reported significant reductions for the APQ Poor Monitoring/Supervision subscale at post-test ($b = -2.00$, 95% CI [-3.23, -0.77], $p = 0.001$) and at follow-up ($b = -1.41$, 95% CI [-2.73, -0.09], $p = 0.036$). In addition, significant reductions were reported for the PS Overreactive parenting subscale at post-test ($b = -0.41$, 95% CI [-

0.71, -0.12], $p = 0.006$), although this affect diminished substantially at follow-up ($b = -0.30$, 95% CI [-0.61, 0.00], $p = 0.051$).

There were significant and sustained reductions in child problem behaviour after the intervention, with caregivers in the intervention group reporting 65% reduced risk of child problem behaviour at post-test (IRR = 0.35, 95% CI [0.28, 0.44], $p < 0.001$) and 70% reduced risk at follow-up (IRR = 0.30, 95% CI [0.23, 0.38], $p < 0.001$). Similarly, child problem behaviour intensity also reduced significantly at post-test ($b = -17.79$, 95% CI [-25.47, -10.10], $p < 0.001$) and at follow-up ($b = -13.48$, 95% CI [-21.26, -5.71], $p < 0.001$). Caregivers in the intervention condition also reported significantly lower overall depression, anxiety, and stress, as measured by the DASS total scale, both at post-test and at follow-up relative to the control group. Specifically, caregivers reported 35% reduced risk of mental health problems at post-test (IRR = 0.65, 95% CI [0.56, 0.75], $p = 0.008$) and 32% reduced risk at follow-up (IRR = 0.68, 95% CI [0.58, 0.80], $p = 0.37$).

Table 16

ITT and CACE Models for Primary Outcomes: Thailand

	ITT analysis			CACE analysis: attend ≥ 4 sessions			CACE analysis: attend ≥ 6 sessions		
Outcome	*b* [95%CI] [a]	IRR [95% CI] [b]	*p*	*b* [95%CI] [a]	IRR [95% CI] [b]	*p*	*b* [95%CI] [a]	IRR [95% CI] [b]	*p*
HOME Harsh parenting [1]									
Posttest	-0.58 [-0.95, -0.20]	0.56 [0.36, 0.86]	.003	-0.66 [-1.10, -0.21]	0.52 [0.31, 0.85]	.003	-0.76 [-1.28, -0.24]	0.47 [0.26, 0.83]	.004
Follow-up	-0.62 [-1.00, -0.24]	0.54 [0.34, 0.82]	.001	-0.72 [-1.16, -0.28]	0.49 [0.29, 0.79]	.001	-0.84 [-1.35, -0.33]	0.43 [0.24, 0.76]	.001
ICAST Overall [2]									
Posttest	-0.97 [-1.42, -0.53]	0.37 [0.32, 0.45]	<.001	-1.17 [-1.69, -0.66]	0.31 [0.25, 0.38]	<.001	-1.33 [-1.92, -0.75]	0.26 [0.21, 0.33]	<.001
Follow-up	-0.88 [-1.37, -0.38]	0.41 [0.34, 0.50]	<.001	-1.03 [-1.60, -0.45]	0.36 [0.29, 0.44]	<.001	-1.18 [-1.83, -0.52]	0.31 [0.24, 0.39]	<.001
ICAST Physical abuse [2]									
Posttest	-0.96 [-1.52, -0.39]	0.38 [0.30, 0.49]	<.001	-1.17 [-1.81, -0.54]	0.31 [0.23, 0.42]	<.001	-1.36 [-2.09, -0.62]	0.26 [0.18, 0.36]	<.001
Follow-up	-0.93 [-1.66, -0.19]	0.40 [0.29, 0.53]	.013	-1.12 [-1.95, -0.28]	0.33 [0.23, 0.47]	.009	-1.29 [-2.25, -0.33]	0.27 [0.18, 0.41]	.008
ICAST Emotional abuse [2]									
Posttest	-0.98 [-1.47, -0.48]	0.38 [0.30, 0.47]	<.001	-1.10 [-1.68, -0.52]	0.33 [0.26, 0.43]	<.001	-1.24 [-1.91, -0.57]	0.29 [0.21, 0.39]	<.001
Follow-up	-0.83 [-1.35, -0.32]	0.44 [0.34, 0.55]	.002	-0.92 [-1.52, -0.32]	0.40 [0.30, 0.52]	.003	-1.04 [-1.74, -0.35]	0.35 [0.26, 0.48]	.003

Note. Poisson logistic regression analyses. [a] Unstandardized beta coefficients computed using robust sandwich estimators; [b] Incidence rate ratio; 95 Cl = 95% confidence intervals. [1]The Home Observation for Measurement of Environment Inventory Abusive Harsh Parenting subscale; [2] ICAST = International Society for the Prevention of Child Abuse and Neglect Child Abuse Screening Tool.

Table 17

ITT and CACE Models for Secondary Outcomes: Thailand

	ITT analysis			CACE analysis: attend ≥ 4 sessions			CACE analysis: attend ≥ 6 sessions		
Outcome	*b* [95%CI] [a]	IRR [95% CI] [b]	*p*	*b* [95%CI] [a]	IRR [95% CI] [b]	*p*	*b* [95%CI] [a]	IRR [95% CI] [b]	*p*
PARYC Total [1]									
Posttest	7.46 [2.92, 11.99]		.001	8.05 [2.76, 13.32]		.002	9.02 [3.13, 14.92]		.002
Follow-up	6.61 [2.31, 10.92]		.003	7.25 [2.23, 12.28]		.003	8.17 [2.46, 13.89]		.003
PS Over-reactivity [2]									
Posttest	-0.41 [-0.71, -0.12]		.006	-0.48 [-0.82, -0.14]		.006	-0.54 [-0.94, -0.15]		.007
Follow-up	-0.30 [-0.61, 0.00]		.051	-0.36 [-0.71, -0.02]		.040	-0.41 [-0.80, -0.01]		.045
APQ Monitoring [3]									
Posttest	-2.00 [-3.23, -0.77]		.001	-2.14 [-3.50, -0.79]		.002	-2.62 [-4.25, -0.98]		.002
Follow-up	-1.41 [-2.73, -0.09]		.036	-1.44 [-2.89, 0.00]		.050	-1.85 [-3.60, -0.11]		.037
ECBI Problem [4]									
Posttest	-1.04 [-1.56, -0.51]	0.35 [0.28, 0.44]	<.001	-1.15 [-1.72, -0.57]	0.32 [0.24, 0.41]	<.001	-1.33 [-2.00, -0.66]	0.26 [0.20, 0.35]	<.001
Follow-up	-1.21 [-1.83, -0.60]	0.30 [0.23, 0.38]	<.001	-1.34 [-2.01, -0.68]	0.26 [0.20, 0.34]	<.001	-1.56 [-2.33, -0.78]	0.21 [0.15, 0.29]	<.001
ECBI Intensity [4]									
Posttest	-17.79 [-25.47, -10.10]		<.001	-20.22 [-28.92, -11.52]		<.001	-22.91 [-32.89, -12.94]		<.001
Follow-up	-13.48 [-21.26, -5.71]		<.001	-15.51 [-24.45, -6.56]		<.001	-17.64 [-28.09, -7.19]		<.001
DASS total [6]									
Posttest	-0.44 [-0.76, -0.11]	0.65 [0.56, 0.75]	.008	-0.47 [-0.83, -0.11]	0.62 [0.53, 0.73]	.010	-0.53 [-0.94, -0.12]	0.59 [0.49, 0.70]	.011
Follow-up	-0.38 [-0.74, -0.02]	0.68 [0.58, 0.80]	.037	-0.40 [-0.80, -0.00]	0.67 [0.56, 0.79]	.049	-0.44 [-0.90, 0.02]	0.64 [0.53, 0.78]	.056

Note. [a] Unstandardized beta coefficient computed using robust sandwich estimators; [b] Incidence rate ratio; 95 Cl = 95% confidence intervals. [1] PARYC = Parenting Young Children Scale; [2] PS = Parenting Scale; [3] APQ = Alabama Parenting Questionnaire; [4] ECBI = Eyberg Child Behaviour Inventory; [5] DASS = Depression, Anxiety, and Stress Scale.

CACE Analyses: Thailand. Two CACE models were examined for each outcome, using the two thresholds of attendance discussed above. The results for the primary and secondary outcomes are reported in the right-hand columns of Table 16 and Table 17, respectively. After controlling for baseline scores, CACE results for harsh parenting showed that estimates of programme effects among caregivers with moderate and high attendance were stronger than those of the ITT analyses, with *p*-values remaining relatively similar. At post-test, there was 44% reduced risk of harsh parenting for all caregivers allocated to the intervention, whereas those with moderate attendance had 48% reduced risk (IRR = 0.52, 95% CI [0.31, 0.85], $p = 0.003$) and high attenders had 53% reduced risk (IRR = 0.47, 95% CI [0.26, 0.83], $p = 0.003$). This pattern was also observed at follow-up, where moderate (IRR = 0.49, 95% CI [0.29, 0.79], $p = 0.001$) and high attending caregivers (IRR = 0.43, 95% CI [0.24, 0.76], $p = 0.001$) had 51% and 57%, respectively, reduced risk of harsh parenting, in comparison to ITT results which indicated a smaller reduced risk of 46%.

For overall child maltreatment, CACE results also showed that programme effects among moderate attenders (post-test: IRR = 0.31, 95% CI [0.25, 0.38], $p = < 0.001$; follow-up: IRR = 0.36, 95% CI [0.29, 0.44], $p = < 0.001$) and high attenders (post-test: IRR = 0.26, 95% CI [0.21, 0.33], $p = < 0.001$; follow-up: IRR = 0.31, 95% CI [0.24, 0.76], $p < 0.001$) were stronger than those of the ITT analyses, with high attendance yielding the greatest reductions in child maltreatment. Caregivers in the entire intervention group reported 63% reduced risk of overall child maltreatment at post-test, moderate attenders reported 69% reduced risk, and high attenders reported 74% reduced risk. Likewise, at follow-up, caregivers in the intervention group reported 59% reduced risk of overall child maltreatment, while moderate attenders reported 64% reduced risk and high attenders reported 69% reduced risk.

This trend was also observed for the physical and emotional abuse subscales. Caregivers with high attendance reported greater reductions in physical abuse (post-test: 0.26 [0.18, 0.36], $p < 0.001$; follow-up: IRR = 0.27, 95% CI [0.18, 0.41], $p = 0.008$) than moderately attending caregivers (post-test: IRR = 0.31, 95% CI [0.23, 0.42], $p < 0.001$; follow-up: IRR = 0.33, 95% CI [0.23, 0.47], $p = 0.009$) or than those in the entire intervention group. Similarly, reductions in reported emotional abuse were smallest in the entire intervention group, followed by moderate attenders (post-test: IRR = 0.33, 95% CI [0.26, 0.43], $p < 0.001$; follow-up: IRR = 0.40, 95% CI [0.30, 0.52], $p = 0.003$), with high attending caregivers reporting the greatest reductions (post-test: IRR = 0.29, 95% CI [0.21, 0.39], $p < 0.001$; follow-up: 0.35 [0.26, 0.48], $p = 0.003$).

Similar patterns were observed for the secondary outcomes, where CACE estimates were both larger than the ITT estimates and increased as the criterion for engagement became more stringent. For positive parenting, CACE models indicated significant programme effect on overall positive parenting for moderate (post-test: $b = 8.05$, 95% CI [2.76, 13.32], $p = 0.002$; follow-up: $b = 7.25$, 95% CI [2.23, 12.28], $p = 0.003$) and high attending caregivers (post-test: $b = 9.02$, 95% CI [3.13, 14.92], $p = 0.002$; follow-up: $b = 8.17$, 95% CI [2.46, 13.89], $p = 0.003$), although p-values remained similar across the ITT and CACE models. Similarly, at post-test, CACE estimates for the PS over-reactive parenting subscale were greater compared to the ITT estimates, with significance levels remaining relatively consistent.

While there was little evidence of a statistically significant effect for caregiver over-reactivity at follow-up in the ITT analyses ($b = -0.30$, 95% CI [-0.61, 0.00], $p = 0.051$), the CACE results indicate significant intervention effects for both moderate ($b = 0.36$, 95% CI [-0.71, -0.02], $p = 0.040$) and high attenders ($b = -0.41$, 95% CI [-0.80, -0.01], $p = 0.045$). At post-test, CACE estimates for the APQ Poor Monitoring/Supervision subscale consistently

increased compared to the ITT estimates, with significance levels remaining similar across models. At follow-up, poor monitoring and supervision only decreased more for high attendance (*b* = -1.85 [-3.60, -0.11], *p* = 0.037) compared to caregivers with moderate attendance (*b* = -1.44, 95% CI [-2.89, 0.00], *p* = 0.050) or those in the entire intervention group.

CACE estimates also indicate greater programme effect than ITT estimates for child problem behaviour and intensity behaviour. Moderately attending caregivers reported 68% less child problem behaviour at post-test (IRR = 0.32, 95% CI [0.24, 0.41], *p* < 0.001) and 74% at follow-up (IRR = 0.26 [0.20, 0.34], *p* < 0.001). Similarly, those with high attendance reported 74% reduced risk of child problem behaviour at post-test (IRR = 0.26, 95% CI [0.20, 0.35], *p* < 0.001) and 79% reductions at follow-up (IRR = 0.21, 95% CI [0.15, 0.29], *p* < 0.001). This pattern was not observed for child problem behaviour intensity, where both moderate (post-test: *b* = -20.22, 95% CI [-28.92, -11.52], *p* < 0.001; follow-up: *b* = -15.51, 95% CI [-24.45, -6.56], *p* < 0.001) and high attending caregivers (post-test: *b* = 22.91, 95% CI [-32.89, -12.94], *p* < 0.001; follow-up: *b* = -17.64, 95% CI [-28.09, -7.19], *p* < 0.001) reported slightly smaller programme effects at follow-up compared to post-test.

Compared to the ITT models, CACE models also indicate consistent estimate increases for the DASS total scale. Caregivers in the entire intervention group reported 35% reduced risk of overall depression, anxiety, and stress at post-test, while those with moderate attendance reported 38% reduced risk (IRR = 0.62, 95% CI [0.53, 0.73], *p* = 0.010) and high attendance reported 41% reduced risk (IRR = 0.59 [0.49, 0.70], *p* = 0.011). Compared to both the entire intervention group and moderate attenders (IRR = 0.67, 95% CI [0.56, 0.79], *p* = 0.049), high attendance yielded greater reductions in caregiver mental health problems at follow-up (IRR = 0.64, 95% CI [0.53, 0.78], *p* = 0.056). Although the *p*-value was slightly greater than the traditional cut-off of .05, IRRs and CIs indicate statistical significance.

Overall, CACE estimates appeared to be larger than ITT estimates. Furthermore, as the attendance thresholds increased from 50% to 75% of the programme, the intervention effects consistently increased. These results suggest that the programme was the most effective for caregivers who attended at least 75% of the intervention. To further test the robustness of the results, the most stringent CACE model was expanded on to test the effects of attendance in addition to baseline predictors of both programme outcome and attendance.

CACE Analyses with Covariates: Thailand. Since attendance was strongly associated to caregiver education and caregiver age (see Table 8), these variables were included as covariates in the model. In addition, models also controlled for child age and child gender to account for potential dependency on outcomes. Results for the adjustment models are presented in Appendix M. When controlling for covariates, the magnitude of programme effects was comparable to the model without covariates. One noticeable difference was that intervention effects for high attendance had a *p*-value smaller than .05 for the DASS total scale at follow-up (IRR = 0.57, 95% CI [0.46, 0.70], $p = 0.012$) which was not seen in the model without the covariates.

Intention-To-Treat Analyses: Philippines. Intervention effects are presented in the left-hand columns of Table 18 (primary outcomes) and Table 19 (secondary outcomes). Controlling for baseline scores, ITT analyses indicated that caregivers in the intervention group reported 48% fewer incidents of overall child maltreatment based on the ICAST total scale at post-test (IRR = 0.52, 95% CI [0.46, 0.58], $p < 0.001$) and 30% fewer at follow-up (IRR = 0.70 [0.62, 0.79], $p = 0.046$) relative to the control group. Analyses also showed significant reductions in physical and emotional abuse subscales at post-test and follow-up based on the IRRs and CIs. Caregivers in the intervention reported 57% less child physical abuse compared to the control group at post-test (IRR = 0.43, 95% CI [0.33, 0.56], $p = 0.002$) and 34% less child physical abuse at follow-up (IRR = 0.66, 95% CI [0.52, 0.83], $p = 0.117$).

Similarly, caregivers in the intervention group reported 43% less child emotional abuse compared to the control group at post-test (IRR = 0.57, 95% CI [0.48, 0.68], $p < 0.001$), with 25% less emotional abuse at follow-up (IRR = 0.75, 95% CI [0.63, 0.89], $p = 0.111$).

In terms of the secondary outcomes, there appear to be no significant intervention effects for overall positive parenting, as measured by the PARYC total scale (post-test: b = 4.01, 95% CI [-0.41, 8.43], $p = 0.075$; follow-up: b =1.76, 95% CI [-2.66, 6.18], $p = 0.436$). Dysfunctional parenting, on the other hand, showed significant decreases for caregivers in the intervention group at post-test (b = -9.64, 95% CI [-13.37, -5.91], $p < 0.001$) but not at follow-up (b = -3.66, 95% CI [-8.65, 1.30], $p = 0.147$).

Based on the IRR CIs overlapping 1.00, it appears as though there were no significant differences between the intervention group and the control group for child problem behaviour at post-test (IRR = 0.90, 95% CI [0.78, 1.03], $p = 0.544$) or follow-up (IRR = 1.21, 95% CI [1.05, 1.39], $p = 0.259$). Non-significant findings were also evident for child problem behaviour intensity at post-test (b =-6.96, 95% CI [-14.31, 0.38], $p = 0.063$) and at follow-up (b =-3.10, 95% CI [-10.50, 4.30], p = 0.411). In addition, results indicate that there was no significant difference between the intervention and the control group in terms of overall caregiver mental health problems, although caregivers appeared to report lower levels of mental health problems at post-test (b = -0.29, 95% CI [-2.70, 2.13], $p = 0.816$) and higher levels of mental health problems at follow-up (b = 0.22, 95% CI [-2.27, 2.71], $p = 0.863$).

Table 18

ITT and CACE Models for Primary Outcomes: Philippines

	ITT analysis			CACE analysis: attend ≥ 6 sessions			CACE analysis: Attend ≥ 9 sessions		
Outcome	*b* [95% CI] [a]	IRR [95%CI] [b]	*p*	*b* [95% CI] [a]	IRR [95%CI] [b]	*p*	*b* [95% CI] [a]	IRR [95%CI] [a]	*p*
ICAST Overall[1]									
Posttest	-0.65 [-0.95, -0.35]	0.52 [0.46, 0.58]	< .001	-0.91 [-1.29, -0.53]	0.40 [0.34, 0.48]	< .001	-1.48 [-2.02, -0.91]	0.23 [0.18, 0.30]	< .001
Follow-up	-0.35 [-0.70, -0.01]	0.70 [0.62, 0.79]	.046	-0.52 [-0.98, -0.05]	0.60 [0.50, 0.70]	.030	-0.86 [-1.60, -0.13]	0.42 [0.33, 0.55]	.022
ICAST Physical abuse[1]									
Posttest	-0.84 [-1.35, -0.32]	0.43 [0.33, 0.56]	.002	-1.13 [-1.80, -0.45]	0.32 [0.22, 0.45]	.001	-1.83 [-2.92, -0.74]	0.16 [0.09, 0.28]	.001
Follow-up	-0.42 [-0.94, 0.10]	0.66 [0.52, 0.83]	.117	-0.60 [-1.30, 0.10]	0.56 [0.40, 0.75]	.094	-0.99 [-2.14, 0.15]	0.37 [0.22, 0.61]	.089
ICAST Emotional abuse[1]									
Posttest	-0.56 [-0.88, -0.24]	0.57 [0.48, 0.68]	< .001	-0.80 [-1.23, -0.38]	0.45 [0.36, 0.56]	< .001	-1.30 [-1.91, -0.69]	0.27 [0.19, 0.39]	< .001
Follow-up	-0.29 [-0.64, 0.07]	0.75 [0.63, 0.89]	.111	-0.41 [-0.89, 0.07]	0.66 [0.53, 0.83]	.095	-0.66 [-1.42, 0.09]	0.51 [0.36, 0.72]	.084

Note. Poisson logistic regression analyses. [a] Unstandardized beta coefficients computed using robust sandwich estimators; [b] Incidence rate ratio; 95 Cl = 95% confidence intervals. [1] International Society for the Prevention of Child Abuse and Neglect Child Abuse Screening Tool.

Table 19

ITT and CACE Models for Secondary Outcomes: Philippines

	ITT analysis			CACE analysis: attend ≥ 6 sessions			CACE analysis: Attend ≥ 9 sessions		
Outcome	*b* [95% CI]	IRR [95% CI]	*p*	*b* [95% CI]	IRR [95% CI]	*p*	*b* [95% CI]	IRR [95% CI]	*p*
PARYC Total [1]									
Posttest	4.01 [-0.41, 8.43]		.075	5.48 [-0.51, 11.48]		.073	9.28 [-0.51, 19.06]		.063
Follow-up	1.76 [-2.66, 6.18]		.436	2.49 [-3.38, 8.36]		.406	4.84 [-4.69, 14.36]		.320
PS Dysfunctional parenting [2]									
Posttest	-9.64 [-13.37, -5.91]		<.001	-13.32 [-18.45, -8.18]		<.001	-20.80 [-29.83, -11.78]		.004
Follow-up	-3.66 [-8.65, 1.30]		0.147	-5.57 [-11.99, 0.85]		.089	-6.39 [-16.83, 4.05]		.230
ECBI Problem [3]									
Posttest	-1.22 [-3.74, 1.31]	0.90 [0.78, 1.03]	.544	-1.55 [-4.92, 1.82]	0.85 [0.71, 1.05]	.5.65	-2.28 [-7.79, 3.22]	0.79 [0.57, 1.10]	.599
Follow-up	0.71 [-1.62, 3.04]	1.21 [1.05, 1.39]	.259	0.97 [-2.10, 4.06]	1.32 [1.08, 1.62]	.2.49	1.79 [-3.25, 6.83]	1.64 [1.16, 2.33]	.236
ECBI Intensity [3]									
Posttest	-6.96 [-14.31, 0.38]		.063	-9.01 [-19.01, 0.98]		.077	-15.23 [-32.12, 1.65]		.077
Follow-up	-3.10 [-10.50, 4.30]		.411	-4.01 [-13.93, 5.91]		.428	-6.84 [-23.72, 10.03]		.427
DASS Total [4]									
Posttest	-0.29 [-2.70, 2.13]		.816	-0.57 [-3.71, 2.57]		.722	-0.81 [-6.10, 4.49]		.765
Follow-up	0.22 [-2.27, 2.71]		.863	0.32 [-2.94, 3.58]		.847	0.50 [-4.98, 5.97]		.859

Note. [a] Unstandardized beta coefficient computed using robust sandwich estimators; [b] Incidence rate ratio; 95 Cl = 95% confidence intervals. [1] Parenting Young Children Scale; [2] Parenting Scale; [3] Eyberg Child Behaviour Inventory; [4] Depression, Anxiety, and Stress Scale.

CACE Analyses: Philippines. The effects of attendance based on CACE estimates are shown along with the ITT results in Table 18 and Table 19. Two CACE models were specified, using the moderate (attended at least 50% or 6 out of 12 sessions) and high (attended at least 75% or 9 out of 12 sessions) thresholds of attendance. Controlling for baseline scores, CACE results for overall child maltreatment showed that estimates of programme effects among moderate and high attending caregivers were stronger than those of the ITT analyses, with high attendance yielding larger effects. CACE results showed 60% reduced risk for moderate attendance (IRR = 0.40, 95% CI [0.34, 0.48], $p < 0.001$) and 77% reduced risk for high attendance at post-test (IRR = 0.23, 95% CI [0.18, 0.30], $p < 0.001$) whereas ITT results showed 48% reductions. Similarly, while ITT results showed 30% reductions in overall child maltreatment at follow-up, CACE estimates indicated 40% reduced risk for moderate attenders (IRR = 0.60, 95% CI [0.50, 0.70], $p = 0.030$) and 58% reduced risk for high attenders (IRR = 0.42, 95% CI [0.33, 0.55], $p = 0.022$).

For physical and emotional abuse, programme effects among moderate and high attending caregivers were also stronger than those of the ITT analyses. At post-test, the reduced risk of physical abuse for moderate (IRR = 0.32, 95% CI [0.22, 0.45], $p = 0.001$) and high attending caregivers (IRR = 0.16 , 95% CI [0.09, 0.28], $p = 0.001$) was 68% and 84%, respectively, which was higher than in the ITT model (57% reduced risk). Similarly, the reduced risk for emotional abuse among moderate attenders was 55% (IRR = 0.45, 95% CI [0.36, 0.56], $p < 0.001$) and 73% for high attenders (IRR = 0.27 [0.19, 0.39], $p < 0.001$), respectively, which was also higher than the ITT model (43% reduced risk). This pattern was also observed at follow-up both for physical (moderate attendance: IRR = 0.56, 95% CI [0.40, 0.75], $p = 0.094$; high attendance: IRR = 0.37, 95% CI [0.22, 0.61], $p = 0.089$) and emotional abuse (moderate attendance: IRR = 0.66, 95% CI [0.53, 0.83], $p = 0.095$; high attendance: IRR = 0.51, 95% CI [0.36, 0.72], $p = 0.084$).

These patterns were also observed for secondary outcomes (see Table 19), although dysfunctional parenting at post-test appeared to be the only outcomes with significant programme effects – also identified in the ITT results. Caregivers with high attendance reported the largest reductions in dysfunctional parenting at post-test (b = -20.80, 95% CI [-29.83, -11.78], p = 0.004), followed by moderate attenders (b = -13.32, 95% CI [-18.45, -8.18], p < 0.001). As seen in the ITT results, programme effects for overall positive parenting were not significant in the CACE models at post-test or at follow-up. In addition, no significant programme effects were reported for child problem behaviour, child problem behaviour intensity, or overall caregiver mental health.

CACE Analyses with Covariates: Philippines. Because it showed the greatest programme effects, the high attendance CACE model was expanded to test the accuracy of the results. Baseline covariates, including caregiver health, emotional abuse, positive parenting, intimate partner violence, and child gender were added as predictors of attendance and outcome because these were identified as the strongest predictors of attendance in the analyses above (see Table 11). In addition, models also controlled for child age to account for potential dependency in outcomes. Results of CACE models with covariates are reported in Appendix N.

Controlling for covariates, programme effects were, for the most part, comparable to the model without covariates, although p-values differed substantially in some instances. Similar to the simple model without covariates, effects were large and significant for overall child maltreatment, physical abuse, and emotional abuse at post-test. However, unlike the simple model, the model with covariates for follow-up assessments also showed significant reductions in physical abuse (IRR = 0.37, 95% CI [0.22, 0.63], p = 0.040) and emotional abuse (IRR = 0.49, 95% CI [0.35, 0.69], p = 0.032) based on both the p-values and the IRR CIs. Similarly, there were also significant reductions in child problem behaviour intensity in

the covariate model (b = -16.48, 95% CI [-32.69, -0.26], p = 0.046), although the wide confidence intervals suggest instability of the results.

Conclusion

To conclude, this chapter investigated factors associated with caregiver attendance in the programmes. It also evaluated programme effects while accounting for varying levels of attendance. Overall, attendance did not appear to be associated with many of the variables identified in the literature. However, CACE results indicate a unique, positive relationship between attendance and participant outcomes. The chapter that follows will discuss these findings in more detail. It will also discuss what implications these findings have for policy, practice, and research, as well as the study's limitations and strengths.

Chapter 4: Discussion

The objective of this present study was to extend knowledge and understanding of the role of attendance in parenting programmes in the understudied settings of LMICs. Specifically, this research set out to examine what factors are associated with attendance and how attendance impacts caregiver and child outcomes. This was undertaken in the context of secondary data analysis of two small-scale RCTs of PLH for Young Children, implemented outside its country of origin (i.e., South Africa), in Thailand and the Philippines.

Summary of Results

Attendance rates

Overall session attendance in Thailand was 82.3%, which is considerably higher than the average rate of 69.8% in the original PLH for Young Children trial in South Africa (Ward et al., 2020). The attendance rate was also higher than the average attendance rates of 72% reported in parenting programmes in HICs (Chacko et al., 2016). One reason for high attendance may be that because of the scarcity of services in Udon Thani Province, caregivers were particularly eager to receive support. Another possibility may be because the programme was delivered by nurses and community health workers within the public health system and not through social services, of whom families are sometimes afraid because it comes with the threat of having their children removed (Featherstone et al., 2014).

Surprisingly, session attendance in the Philippines was lower at 61.8%, despite being part of a conditional cash transfer system. Reasons for this may be that the amount provided in the conditional cash transfer was not sufficient motivation to attend sessions. Although not part of a conditional cash transfers, some studies in LMICs report even lower attendance rates, such as 50% for the 14-session group-based PLH for Adolescents programme trial which was integrated within the Isibindi programme, a community-based protection intervention in South Africa (Cluver et al., 2018). In HICs, low attendance rates are also not

uncommon for parenting interventions even when delivered within existing services. For instance, a study of a group-based Triple P parenting programme delivered though the public system in Birmingham, United Kingdom, reported an average attendance rate of 40% (Little et al., 2012). The differing attendance rates between Thailand and the Philippines point to the need for further research on the conditions required to successfully implement programmes within different types of routine delivery systems.

While qualitative interviews were not conducted in Thailand, interviews conducted by the research team with participants in the Philippines revealed that there were several major barriers to attendance, including schedule conflicts despite group sessions taking place on Saturdays. Even though approximately one-third of caregivers were unemployed, many had informal work such as selling products at markets on weekends. Participants also noted that attending church activities, looking after children, and finishing chores on weekends prevented their attendance. These barriers point to the need to find creative ways to reach caregivers more effectively. Providing childcare may be one way to help remove barriers to attendance (Axford et al., 2012). Childcare was provided in Thailand but not in the Philippines, perhaps explaining the differing attendance rates between the two studies.

Workplace delivery or ensuring that sessions take place outside of working hours may be another way of making programmes more accessible (Sanders et al., 2011). Alternative approaches may also include delivery via digital formats which could reduce multiple logistical barriers common in-person programmes (Baumel et al., 2016; Breitenstein et al., 2014). Digital delivery may also provide a more scalable and cost-effective approach to reaching large numbers of families in need of support (Hall & Bierman, 2015). As of May 2020, internet penetration (i.e., the percentage of the total population that use the internet) was found to be 81.7% in Thailand and 71.1% in the Philippines (Moore, 2020), suggesting that digital delivery may be particularly promising in these contexts.

Predictors of Attendance

To identify barriers and facilitators of attendance in PLH for Young Children in Thailand and the Philippines, the current study took an exploratory approach and tested a range of factors that have previously been linked to attendance in parenting programmes. These factors were related to household economic status and caregiver education level, caregiver social and health wellbeing, caregiver and child behaviour, and sociodemographic characteristics.

As in many previous studies of parenting programmes, socio-economic status was not associated with attendance in either of the RCTs. However, in Thailand, findings indicate that caregiver's education level influenced their attendance in the programme. Specifically, caregivers with lower levels of education attended more sessions. These findings are consistent with the PLH for Young Children trial in South Africa which found some indication (although not statistically significant) that caregivers who had completed high school were more at risk of missing a session (Wessels, 2017). A possible explanation for these finding is that employment among caregivers with higher levels of education may have been more likely and they were therefore less available to attend sessions. In contrast to these findings, most previous studies of parenting programmes in HICs indicate that lower education level is a risk factor for poor attendance and more dropout (Laxman et al., 2019; Reyno & McGrath, 2006). There may therefore be a dynamic in LMICs where caregivers want to invest in their children's wellbeing to give them the chances they did not have growing up, which may be absent in HICs.

Similar to other studies, caregiver mental health was not significantly associated with attendance. On the other hand, less healthy caregivers, as measured by the Medical Outcomes Study Short Form-12 Health Survey, were significantly more likely to attend in the Philippines. Although this tool assesses general health and does not make a distinction

between physical and mental health, responses by caregivers may well have captured some aspects of mental health. In addition, caregiver health was positively correlated to number of household assets, suggesting that poorer health may be linked to lower socio-economic status. It is therefore possible that caregivers with poorer health, and hence lower economic status, were more incentivised to attend in the Philippines so that they could receive the cash grant. Since caregiver health has very rarely been examined in the context of parenting programmes, there is a need to examine this potential facilitator of attendance further.

Overall, findings that lower education and poorer health led to increased attendance are encouraging as they suggest that the programmes can successfully reach vulnerable families in LMICs. However, results also showed that caregivers who experienced more intimate partner violence were more at risk of poor attendance in the Philippines. These findings are inconsistent with findings from the PLH for Young Children and the PLH for Adolescents trials in South Africa, which indicated no association between intimate partner violence exposure and session attendance (Shenderovich et al., 2018; Wessels, 2017). Social and health barriers such as intimate partner violence may very well inhibit attendance and need to be further investigated so that families facing such adversities are reached during implementation. It may be that partners of women in the group forbade them to attend, threatening violence if they did; or that attempts to try the new skills at home provoked conflict; or that the aftermath of violent episodes caused them shame (for instance, bruises) or increased the need for their presence with their children (for instance, to provide comfort or protection). Future studies would benefit from examining the role of partner coercion on attendance in parenting programmes.

Although some studies have found that higher levels of child behavioural problems were associated with higher attendance (e.g. Heinrichs et al., 2005), other studies have found no such relationship (Shenderovich et al., 2018; Wessels, 2017). Similarly, child behaviour

was not related to caregiver attendance in both RCTs. However, some parenting dimensions did appear to be related to attendance. In Thailand, caregivers who used more proactive parenting strategies were less at risk of missing a session, although the size of effect was very small. This finding is consistent with the PLH for Adolescents trial which indicated that higher levels of positive and involved parenting was associated with higher caregiver attendance (Shenderovich et al., 2018). One reason for this finding may be that caregivers who were more proactive were more interested in learning new parenting techniques and therefore attended more sessions, or more able to organise childcare, chores, and work-related responsibilities so that they were free to attend.

In the Philippines, caregivers who used higher levels of emotional abuse were slightly more likely to attend sessions. This finding aligns with the PLH for Young Children trial in South Africa which found that caregivers who used higher levels of psychological discipline were at lower risk of missing a session (Wessels, 2017). A reason for these findings may be that caregivers found learning positive alternative to negative parenting practices helpful and were thus motivated to continue attending (Wessels, 2017).

Several previous studies have found that younger parental age was linked to more attrition and less attendance (e.g. Reyno & McGrath, 2006). While caregiver age was not associated with attendance in the Philippines, older caregivers were significantly more likely to attend sessions in Thailand. Although the mechanism responsible for this relationship in parenting programmes is not clear, some authors speculate that younger caregivers may experience higher levels of parenting stress, which, in turn, affects their ability to participate (Kazdin et al., 1993). In the current study, caregiver age and caregiver education were highly negatively correlated, suggesting that older caregivers were also less educated. Thailand has gone through a dramatic period of economic development and has immensely expanded access to education over the past few decades which would explain why older caregivers

were also less educated. Some studies have also found that higher personal income is associated with less attendance in parenting programmes (Laxman et al., 2019). Therefore, another possible explanation for younger caregivers attending fewer sessions may be that they were also higher earners (due to their higher levels of education) and had less time to attend sessions. It may also be that younger caregivers felt uncomfortable participating in the presence of older ones due to issues of respect and privacy. Additional research is needed to identify why younger caregivers in Thailand were at greater risk of poor attendance and how to support their attendance in future.

Child age was not associated with attendance in either country. However, there was a significant association between caregivers with boys and higher attendance in the Philippines. This finding is inconsistent with the PLH for Young Children trial in South Africa which found a weak, although nonsignificant, association between being a parent of a boy and greater risk of missing a session (Wessels, 2017). A possible reason for this finding in the Philippines study is that caregivers perceived parenting boys as more challenging and therefore felt like they needed extra support. However, very little research has explored the relationship between child demographics such as age and gender and caregiver attendance and thus requires further attention.

In summary, findings suggest that more disadvantaged families participated at similar or higher rates than families with more resources. One interpretation for these findings is that more disadvantaged families may have a greater perceived need for material and psychosocial support and therefore attended at higher rates (Shenderovich et al., 2018). However, few of the factors identified in the existing literature were significantly associated with attendance in the present study. This relatively low number of significant factors is not surprising for two reasons. First, the influence of many of the factors identified from past research were inconsistent (e.g. socio-economic status, marital status, children's behaviour),

studied in the context of HICs, or had rarely been studied in parenting programmes (e.g. caregiver health and intimate partner violence). Second, the programmes examined in the current study had low dropout rates (5%) and, in Thailand, relatively high attendance. Thus, limited variation in the dosage of the programme may have affected the ability to detect associations.

Impact of Attendance on Intervention Outcomes

The present study also examined whether variation in attendance resulted in detectable differences in participants outcomes over time. In previous research, attendance and related implementation indicators such as engagement have not always impacted programme outcomes (Nix et al., 2009; Shenderovich et al., 2019). However, the PLH for Young Children trial in South Africa found a dose-effect relationship at post-assessment, with higher session attendance associated with significant improvements in some participant outcomes (Wessels, 2017). While most dose-response relationships are estimated using intervention arm data only, or by removing noncompliers from the analyses, this study used the novel CACE approach which accounts for confounding effects of moderators while simultaneously maintaining randomisation.

Findings demonstrate that PLH for Young Children Thailand had significant effects on most caregiver and child outcomes, both at the one-month post-test and the three-months post-intervention follow-up. Results also indicate that relative to the ITT analyses, the CACE approach yielded stronger programme effects among caregivers who attended more sessions. For observed harsh parenting, caregivers on average had greater decreases in harsh parenting practices among moderate and high engaging caregivers than in the entire intervention condition, with the strongest and highest risk reductions found at follow-up among caregivers who attended at least 75% of the programme. For overall child maltreatment as well as for physical and emotional abuse, moderate and high attenders had greatest reductions at post-

intervention, with strongest effects found among caregivers who attended at least 75% of the programme. Overall, these findings suggest that improvements in observed harsh parenting and child maltreatment can result from less than full attendance at programme sessions. However, improvements were also larger and stronger among caregivers who attended at least six out of the eight intervention sessions.

Furthermore, the effects of PLH for Young Children Thailand on positive parenting, over-reactive parenting, and poor monitoring and supervision were also larger in magnitude among high attenders at post-test and at follow-up when compared to both moderate attenders and the programme effects identified in the corresponding ITT analyses. Similarly, programme effects for child problem behaviour and child problem behaviour intensity improved when accounting for attendance, with programme effects increasing as the number of sessions attended increased.

The estimation of programme impacts on caregiver metal health problems while accounting for attendance revealed a slightly different pattern in Thailand. While CACE models indicate greater programme effects relative to ITT models, significance levels based on *p*-values decreased with more attendance and were no longer significant for participants who attended at least 75% of the programme at follow-up. When taking covariates into account, the CACE model revealed a significant effect of attendance among caregivers who attended at least 75% of the programme. In the covariate model, higher education levels and having a boy child also predicted significant decreases in caregiver mental health at follow-up. Although child gender was not associated with attendance in Thailand, these findings suggest that caregivers with lower education and girl children may be important subgroups to identify at the start of the programme as they may need additional support to boost their outcomes.

Overall, similar trends were found for primary outcomes in the Philippines. When ITT models were used, programme effects were identified for reductions in overall child maltreatment, physical abuse, and emotional abuse, both at immediate post-intervention and at one-year follow-up. However, CACE estimates revealed that intervention effects were substantially greater among caregivers who attended more sessions, with greatest reductions in reported child maltreatment found among caregivers who attended at least nine out of 12 sessions.

In terms of secondary outcomes, results varied slightly more in the Philippines than in Thailand. The ITT results showed significant reductions in overall dysfunctional parenting, but only immediately post-intervention. Despite CACE analyses revealing no significant moderating effect of attendance at follow-up for this outcome, intervention effects were substantially greater among caregivers who attended more sessions both at immediate post-test and at one-year post-intervention follow-up.

The ITT results demonstrate that the programme had no significant effect on positive parenting, child problem behaviour and child problem behaviour intensity, and caregiver mental health problems, either immediately at post-intervention, or at one-year post-intervention follow-up. Similarly, CACE analyses revealed no significant moderating effect of attendance at post-test or at follow-up although intervention effects were greater among caregivers who attended more sessions. However, when intimate partner violence was included as a covariate, it was associated with significant increases of child problem behaviour intensity at follow-up. CACE estimates accounting for covariates also revealed significant reductions in child problem behaviour intensity immediately post-intervention among caregivers who attended at least 75% of the programme. This finding suggest that caregivers who reported higher rates of child problem behaviour intensity and who experienced co-occurring intimate partner violence may be an important subgroup to identify

at programme intake so that additional support might be provided to boost their outcomes. Future research should examine the moderating impact of intimate partner violence on child behaviour problems.

To summarise, much of the current evidence of the effectiveness of parenting programmes relies on ITT analyses. Overall programme effects are important for understanding whether parenting programmes can work in community settings such as those in Thailand and the Philippines. However, estimating variation in attendance and its impacts on participant outcomes using CACE can help unpack for whom and under what circumstances programmes are effective (Schochet et al., 2014). This is especially helpful for informing recruitment and retention strategies for evidence-based interventions. Overall, ITT results suggest that improvements in targeted outcomes can result from less than full attendance at programme sessions. However, similar to other applications of CACE in prevention interventions (e.g. Connell, 2009; Huang et al., 2014; Stuart et al., 2008), greater CACE effects were observed as the definition of engagement became more stringent. Specifically, CACE models suggest that improvements are larger and stronger among caregivers who attended at least 75% of the programme. These findings highlight the importance of tailoring parenting programmes to families in a way that motivates their attendance as it appears to positively impact outcomes.

Implications for Policy and Practice

There are several important implications for policy and practise that follow from investigating factors associated with attendance as well as from the application of CACE methods in parenting programmes. As PLH for Young Children is increasingly being transported from trial settings to routine service delivery settings, a more nuanced understanding of the factors associated with attendance may inform recruitment and retention strategies used by implementing organisations in ordinary service settings. In addition,

knowing what programme effects can reasonably be expected from specific rates of attendance can help implementation organisations devise support structures and strategies to successfully achieve these attendance rates. Important policy lessons can also be drawn when comparing ITT effects to CACE effects. CACE estimates can disentangle programme effects by distinguishing whether weak or null effects found in ITT analyses are due to small effects among high attenders or due to low rates of attendance in the programme. Specific strategies that could be considered to encourage attendance in parenting programmes, including tailoring interventions for vulnerable subgroups, and integrating programmes within existing services, are discussed below.

Multi-layered System of Support

In the high-stress and low-income settings of LMICs, families face multiple and intersecting challenges characterised by poverty, violence, and illness. Since these contexts often lack access to basic social and health service (Jordans et al., 2016), families may require several different types of support to reduce violence against children. For instance, the WHO INSPIRE framework identifies and recommends seven different evidence-based strategies to combat violence against children including: implementation and enforcement of laws; norms and values, safe environments; parent and caregiver support; income and economic strengthening; response and support services; and education and life skills (WHO, 2016). Nevertheless, there is also a growing evidence-base that programmes which address violence against children, such as parenting interventions, may also act as accelerators on a number of other social and health outcomes. For example, a prospective cohort study of 1176 adolescents 10-19 years living in South Africa (Cluver et al., 2019) found that parenting support was not only associated with no emotional or physical abuse (odds ratio = 2.38, 95% CI [1.65, 3.76]), but also with good mental health (2.13, 95% CI [1.43, 3.15]), no high-risk sex (2.44, 95% CI [1.45, 5.03]), no violence perpetration (2.59, 95% CI [1.63, 4.59]), and no

community violence (2.43, 95% CI [1.65, 3.86]). Additional support could also be provided either through “add-on” components or through multi-layered care packages (Inter-Agency Standing Committee (IASC), 2007). For instance, the PLH for Adolescents programme has recently been adapted by Clowns Without Borders South Africa to include components on HIV/AIDS and sexual health to support families facing this particular challenge. In the context of the current study, it may also be important to consider add-on components for intimate partner violence.

This study found that caregivers who experienced heightened rates of intimate partner violence in the Philippines were more at risk of missing programme sessions and may thus have been less likely to benefit from the programme. Intimate partner violence has been linked to harsh parenting, reduced rates of engagement with children, lower positive parenting as well as child neglect (Chiesa et al., 2018). However, a recent cross-sectional study from Demographic and Health Surveys of over 15,000 households in LMICs found that a direct association between intimate partner violence and child wellbeing was mediated through positive parenting behaviours (Jeong et al., 2020). Indeed, a recent systematic review of 18 studies (eight of which were from LMICs) on the effects of parenting interventions on intimate partner violence suggests that parenting programmes may reduce both overall intimate partner violence and physical female victimisation (Schafer, 2020). Families experiencing intimate partner violence may therefore have the greatest potential to benefit from attending parenting programmes yet face significant barriers to participation. This highlights the importance of providing additional support for this subgroup, possibly by including modules that specifically address partner relationships or by engaging other caregivers such as fathers in the programme.

Involving Other Family Members

Although involving other family members, such as fathers, may reduce programme reach due to limitations on group size, it may lead to improved participation of primary caregivers (Lundahl et al., 2008) and may also support long-lasting behaviour change within families (Errázuriz et al., 2016). There have thus been increasing efforts to engage fathers in parenting programme; however, recruitment rates tend to be low. For instance, the recruitment rate was 5% in the PLH for Adolescents trial in South Africa (Shenderovich et al., 2018). Similarly, low rates of 3% and 14% were achieved in Triple P programmes in Panama and China, respectively (Guo et al., 2016; Mejia et al., 2015). Like previous studies, the programme comprised mostly of female caregivers in Thailand and no male caregivers were included in the Philippines. While the present study found no significant difference between male and female attendance in Thailand, other studies have found that men were at greater risk of poor attendance (Panter-Brick et al., 2014; Shenderovich et al., 2018). Hence, it is important to develop strategies to recruit more male caregivers, but also to understand how programme content can be optimised for fathers and other family members.

Integrating Interventions Within Existing Services

Integrating parenting interventions within existing social and health services may be another avenue to provide additional support to vulnerable families. For example, in addition to the parenting programme, the recipients of the 4Ps programme in the Philippines had access to health and education services as well as financial cash grants. By targeting multiple adversities simultaneously, families may therefore have benefited more than if they had only received a stand-alone intervention like in South Africa. However, with no Filipino counterfactual to compare findings with, the added benefits of conditional cash transfer systems remain unknown.

In addition to offering multiple services, benefits of integrating interventions within existing services may include leveraging existing relationships between families and service providers (Mytton et al., 2014). Accordingly, delivering parenting programmes within established service delivery systems may have the potential to increase participation significantly during wider scale roll-out. However, integrating services with existing ones may come at a steep price because the social and health professionals within these systems are often overworked in most LMICs (Walker et al., 2018). Adding the additional workload of delivering a new intervention on top of the existing services may compromise the quality of delivery. In addition, the programme may not align with the pre-existing principles in the implementing organisation. For instance, parenting programmes such as PLH used participatory approaches for delivery while social care staff may be more familiar with didactic teaching approaches (Doubt et al., 2018). It is therefore important that programme developers work closely with implementing partners to ensure adequate training and programme fit within the delivery system.

Implications for Theory and Research

In addition to policy and practice implications, this study adds to the scarce body of literature on factors associated with attendance in parenting programmes in LMICs. Furthermore, given that this is the first known study to explore attendance variability through CACE in the context of parenting programmes in LMICs, findings add new and rigorous evidence that programme outcomes are positively impacted by higher participant attendance. This study has also identified several additional areas that require further research. The recommendations for such research are discussed below.

Implementation in Routine Settings

The PLH programmes are currently undergoing a wide-spread scale-up in over 25 LMICs (Shenderovich et al., in submission). It is therefore important to conduct research in

settings that are typical for routine services. Delivering programmes outside the controlled conditions of trials is complex and challenging, and often results in loss of effectiveness (Welsh et al., 2010). These moderators of effectiveness include a range of implementation factors, provider/system characteristics, and target population characteristics (Ogden & Fixsen, 2014). Challenges related to implementation factors involve a host of issues such as insufficient resources and infrastructure needed for training, supervision, and technical support (Dodge, 2001; Elliott & Mihalic, 2004). Ensuring implementation fidelity - the degree to which programmes have been implemented as intended (Fixsen et al., 2005), may also be difficult due to demands for local adaptations of programmes, which may, in turn, lead to loss of effects (Ogden & Fixsen, 2014).

Barriers to service providers and systems adopting the programme include variation in commitment and motivation from programme staff, diverse backgrounds in training, expertise among facilitators delivering the programme (Welsh et al., 2010). Other challenges include the need for a local programme champion/coordinator within the organisation who has sufficient persuasion skills and influence to ensure adoption and programme maintenance (Hutchings, Bywater, & Daley, 2007). Finally, moving from relatively homogeneous populations in trials to different community contexts may also result in greater variations in the motivation of families to enrol and attend, and more comorbidity due to less stringent participation criteria (Welsh et al., 2010). Since all these factors may affect attendance and therefore outcomes, it would be beneficial for future research to draw on monitoring and evaluation data from routine service delivery. In doing so, researchers could expand the available pool of data to examine relationships between implementation factors, such as attendance, and family outcomes in real world contexts.

Impact of Attendance Boosters

Although PLH for Young Children has demonstrated promising effects on reducing child maltreatment when integrated within existing service delivery systems in Thailand and the Philippines, additional research using innovative approaches such as factorial experimental designs could help optimise the programmes for scalability (Collins, 2018). Like many parenting interventions (see Axford et al., 2012), the two RCTs used various incentives, or attendance boosters, such as transport reimbursements, financial incentives, meals, text messages, and phone calls to increase participation. However, little is known about the effectiveness and cost-effectiveness of these methods. This is especially important in the context of real-world delivery when implementing organisations typically operate under financial and human resource constraints. It could therefore be beneficial to test differential effects of the intervention package, including attendance boosters and add-on modules focused on retaining at-risk subgroups of caregivers such as those affected by intimate partner violence. This approach could also provide insight into how much training and supervision is necessary to improve participant outcomes at scale.

Recently, PLH for Young Children has been tested in a factorial cluster RCT in Romania, the Republic of Moldova, and North Macedonia (Lachman et al., 2019) to evaluate the effectiveness and cost-effectiveness of three different components of the intervention: engagement (basic versus and enhanced engagement packages), length (5 sessions versus 10 sessions), and fidelity (supervision for facilitators on demand versus structed supervision). Although results are not yet available, findings will be especially beneficial for implementation organisations who want to know how best to improve cost-benefit ratios.

Analytic Approach

To my knowledge, this is the first study to apply the CACE analytic approach to parenting programmes in LMICs. However, several aspects of this approach could be refined

in future research. CACE methods such as the one utilised in this study require a binary attendance variable. However, partial attendance is common in parenting programme which may result in violations of the exclusion restriction assumption. That is, CACE approaches assume that noncompliers in the intervention group (for this study, participants who attended below 50% or 75%) do not benefit from the programme, when in fact they might. In the current study, this was partially addressed in two different ways.

First, since the exclusion restriction assumption depends heavily on where the compliance threshold is set, two different thresholds were set to determine whether there was a substantial difference between caregivers who attended 50% versus those who attended 75%. Indeed, findings indicate that higher attenders benefited more, although it was evident that participants who attended less than 75% also benefited - suggesting potential violations of the assumption. Other studies have set three or even more thresholds to achieve this end (see Connell, 2009). However, given the novelty of applying CACE methods to evaluations of parenting programmes, selecting appropriate and accurate thresholds for "dosage" is challenging. Thus, this study based its definitions on previous literature of prevention and education trials. Additional research could greatly benefit from a collaboration between programme developers, implementers, and statisticians to help determine what could reasonably constitute adequate attendance in parenting programmes such as PLH for Young Children. Identifying optimal definitions of "dosage" is particularly valuable for obtaining robust CACE estimations of outcomes as well as for making comparisons across parenting intervention studies (Huang et al., 2014).

Second, identifying good predictors of attendance has been recommended to reduce susceptibility of violating the reclusion restriction assumption (Frangakis & Rubin, 1999; Jo, 2002). Thus, a third CACE model was tested which included predictors of attendance (identified in the first stage of the analyses) as covariates. CACE covariate models

specifically predicted programme effects for participants who attended at least 75% of the programme. However, predictors of attending at least 75% of the programme may not have been the same as predictors for the entire intervention arm. Future studies should investigate predictors of attendance in accordance with the compliance thresholds used in the CACE models. Another way to further enhance covariate models is to ensure that future studies measure a range of baseline characteristic that may influence attendance rates, including variables that are not commonly measured in parenting programmes. For instance, these might include how long it takes participants to get to the programme site, whether or not they need support with childcare during sessions, and parental motivation to attend sessions. Essentially, the more accurate the baseline variables are at predicting attendance, the more accurate the CACE estimates will be (Stuart et al., 2008). To enhance model accuracy, future studies could therefore benefit greatly from measuring logistical factors as well as parental motivation.

Alternative approaches which prevent violations of the exclusion restriction assumption could also be considered. These might include propensity score matching CACE approaches which could be used to examine the impacts of attendance on outcomes for subgroups of caregivers, where participants in the intervention group with the same attendance rates are matched to participants in the control group based on similar baseline characteristics (Stuart et al., 2008). However, this approach requires a larger sample than was available in the present study. Alternatively, instrumental variable approaches also allow for continuous measures of attendance, which permits more fine-tuned investigations of the relationship between attendance and participant outcomes (Connell, 2009).

In summary, although the current study adds important findings the literature on attendance in parenting interventions in LMICs, future studies could advance this field of research by conducting evaluations on the factors that influence attendance in routine service

settings outside the conditions of well-resourced RCTs. Future research should also investigate the effectiveness and cost-effectiveness of intervention components such as incentives and retention strategies for vulnerable subgroups. Further, although rigorous CACE methods were used to account for attendance variability within the samples, a close collaboration between developers, implementers, and statisticians to identify adequate attendance in parenting programmes could improve the accuracy of results. Additionally, good measures of factors associated with attendance should be collected and alternatives to CACE could be explored if binary indicators of attendance are not desirable.

Limitations and Strengths

This study has several limitations that warrant mentioning. First, observational assessments were only used in Thailand, and this was only for the primary outcome. Since self-report methods may illicit social desirability and thus result in reporting bias (van de Mortel, 2008), future research could benefit from more extensive observational assessments to supplement self-reports. Further, it is also worth noting the potential limitations of solely using self-reports from caregivers. Previous research has shown substantial disagreement between caregiver and child reports, especially in terms of assessing child maltreatment (Sierau et al., 2017) and child mental health (De Los Reyes et al., 2015; Kim et al., 2020). It may therefore be useful to compare the ratings between caregiver self-reports, child self-reports and even proxy-reports from other family and household members in future.

Second, attendance was measured using a composite score from both group sessions and home visits in Thailand (no home visits were offered in the Philippines). However, factors influencing attendance at home visits may differ substantially from those at group sessions. To strengthen retention strategies going forward, research could investigate in what ways baseline characteristics are associated with group sessions versus home visits. Third, although the outcome measures had been used in the Thai and the Filipino pilot studies, none

of the measures has been formally validated in these contexts. Since unique linguistic, contextual, and cultural factors may influence whether these measures measure the same constructs as in the original validation studies, future evaluations could benefit from validation studies in Thailand and the Philippines. Fourth, although effects appeared to have sustained at follow-up for all outcomes in Thailand, follow-up assessment occurred shortly after the programme (3 months), and so it remains unknown whether the reported changes will endure over time. On the other hand, results at one-year follow-up only endured for some of the outcomes in the Philippines. Longer-term follow-up may be needed to establish the degree of intervention maintenance in each context as well as whether there are any delayed intervention effects.

Fifth, sample size calculations were based on the studies' main effect analyses, and it is therefore possible that the analyses of the present study were underpowered. This may have been a particular issue for CACE modelling as missing data were not accounted for. This leads to a sixth limitation. For the CACE analyses, this study ignored potential biases introduced by removing cases that dropped out of the study (not necessarily out of the programme). Although much of the current research on CACE analysis relies on listwise deletion of missing cases (Berg et al., 2017; Jo et al., 2008), full maximum likelihood methods under the assumption that data are missing at random have been recommend (Stuart et al., 2008). However, one complication encountered when there is both missing data and noncompliance/poor attendance is that missingness may be associated with compliance status (Knox et al., 2014; Yau & Little, 2001). As a result, missing cases may not be missing at random and thus may not be appropriately accounted for in maximin likelihood methods. More sophisticated modelling techniques such as Bayesian CACE methods can also account for missingness and should be investigated for their utility in this kind of study. Another limitation of basic CACE methods more generally is that many do not account for clustering

of individuals within groups. Therefore, for this study, potential clustering of caregivers within parenting groups was ignored which may have had biased effects on the results. However, it should be noted that previous analyses conducted by the research team did not find significant intraclass correlation coefficients within assigned parenting groups either in Thailand or the Philippines.

Despite these limitations, this study has numerous strengths. Both trials were conducted in community settings and were delivered as part of existing services. Therefore, factors found to influence attendance are likely similar to those in routine delivery settings, which suggests high external validity of results. That is, due to the nature of delivery within existing services, results may generalise to other families served by these and similar services systems in Thailand and the Philippines. Measures used in the analyses also showed high internal consistency, indicating reliability of findings within the sample populations. Further, this study also provides a rigorous examination of the impacts of PLH for Young Children both at post-intervention and at follow-up while robustly accounting for moderating effects of attendance via CACE analysis. In doing so, analyses upheld randomisation and accounted for potential confounding effects thereby reducing biases commonly experienced in dose-effect modelling. Finally, this study was the first to conduct an in-depth evaluation of attendance in parenting programmes in Southeast Asian contexts such as Thailand and the Philippines and so provides utility for informing retention strategies for vulnerable families in future implementation efforts.

Conclusion

As evidence-based parenting programmes such as PLH for Young Children are increasingly being disseminated in many LMICs, it is important to investigate factors that affect family's attendance in these programmes. It is equally important to determine whether attendance is associated with targeted outcomes and how much "dosage" is required to

achieve sustained effects. This secondary data analysis of two RCTs of PLH for Young Children implemented in the understudied contexts of Thailand and the Philippines makes important contributions to understanding what factors affect attendance and how attendance relates to programme outcomes within existing service delivery systems in LMICs. The promising results suggest that poverty and related disadvantages generally did not inhibit caregivers in their ability to attend the programmes. Nevertheless, exposure to intimate partner violence was an important barrier to attendance in the Philippines, underscoring the need for devising additional support structures for families who face these and other adversities. Furthermore, findings suggest that while the programmes can produce significant improvements in outcomes, the magnitude of these improvements is contingent upon attendance. This further highlights the importance of tailoring parenting programmes to families in a way that motivates their attendance.

Although many research questions remain unanswered, this study makes an important step towards uncovering how families in resource-constrained community settings participate in parenting programmes. The next step is to collaborate with those involved in scaling up parenting programme to further design and test effective and cost-effective strategies to support attendance to ensure that programmes do prevent violence against children and improve child wellbeing outside the controlled conditions of trials.

Reference

Alampay, L., Lachman, J., Landoy, B., Madrid, B., Ward, C., Hutchings, J., Alinea, M., & Gardner, F. (2018). Preventing child maltreatment in low- and middle-income countries: Parenting for Lifelong Health in the Philippines. In S. Verma & A. C. Petersen (Eds.), *Developmental Science and Sustainable Development Goals for Children and Youth* (pp. 277-293). Springer International Publishing. https://doi.org/10.1007/978-3-319-96592-5_15

Anda, R., Felitti, V., Bremner, J., Walker, J., Whitfield, C., Perry, B., Dube, S., & Giles, W. H. (2006). The enduring effects of abuse and related adverse experiences in childhood. *European Archives of Psychiatry and Clinical Neuroscience, 256*(3), 174-186. https://doi.org/10.1007/s00406-005-0624-4

Angrist, J., Imbens, G., & Rubin, D. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association, 91*(434), 444-455. https://doi.org/10.1080/01621459.1996.10476902

Arnold, D., O'Leary, S., Wolff, L., & Acker, M. (1993). The Parenting Scale: A measure of dysfunctional parenting in discipline situations. *Psychological Assessment, 5*(2), 137-144. https://doi.org/10.1037/1040-3590.5.2.137

Ashworth, E., Panayiotou, M., Humphrey, N., & Hennessey, A. (2020). Game on - complier average causal effect estimation reveals sleeper effects on academic attainment in a randomized trial of the Good Behavior Game. *Prevention Science*, *21*(2), 222-233. https://doi.org/10.1007/s11121-019-01074-6

Aspland, H., & Gardner, F. (2003). Observational measures of parent-child interaction: An introductory review. *Child and Adolescent Mental Health, 8*(3), 136-143. https://doi.org/10.1111/1475-3588.00061

Axford, N., Lehtonen, M., Kaoukji, D., Tobin, K., & Berry, V. (2012). Engaging parents in parenting programs: Lessons from research and practice. *Children and Youth Services Review, 34*(10), 2061-2071. https://doi.org/10.1016/j.childyouth.2012.06.011

Baker, C., Arnold, D., & Meagher, S. (2011). Enrollment and attendance in a parent training prevention program for conduct problems. *Prevention Science, 12*(2), 126-138. https://doi.org/10.1007/s11121-010-0187-0

Barlow, J., & Coren, E. (2018). The effectiveness of parenting programs. *Research on Social Work Practice, 28*(1), 99-102. https://doi.org/10.1177/1049731517725184

Barlow, J., Johnston, I., Kendrick, D., Polnay, L., & Stewart-Brown, S. (2006). Individual and group-based parenting programmes for the treatment of physical child abuse and neglect. *Cochrane Database of Systematic Reviews, 3*. https://doi.org/10.1002/14651858.cd005463

Baumel, A., Pawar, A., Kane, J., & Correll, C. (2016). Digital parent training for children with disruptive behaviors: Systematic review and meta-naalysis of randomized trials. *Journal of Child and Adolescent Psychopharmacology, 26*(8), 740-749. https://doi.org/10.1089/cap.2016.0048

Baydar, N., Reid, M., & Webster-Stratton, C. (2003). The role of mental health factors and program engagement in the effectiveness of a preventive parenting program for Head Start mothers. *Child Development, 74*(5), 1433-1453. https://doi.org/10.1111/1467-8624.00616

Bendel, R., & Afifi, A. (1977). Comparison of stopping rules in forward “stepwise” regression. *Journal of the American Statistical Association, 72*(357), 46-53. https://doi.org/10.1080/01621459.1977.10479905

Berg, J., Bradshaw, C., Jo, B., & Ialongo, N. (2017). Using complier average causal effect estimation to determine the impacts of the Good Behavior Game Preventive

intervention on teacher implementers. *Administration and Policy in Mental Health and Mental Health Services Research, 44*(4), 558-571. https://doi.org/10.1007/s10488-016-0738-1

Berkel, C., Mauricio, A., Schoenfelder, E., & Sandler, I. (2011). Putting the pieces together: An integrated model of program implementation. *Prevention Science, 12*(1), 23-33. https://doi.org/10.1007/s11121-010-0186-1

Berkel, C., Sandler, I., Wolchik, S., Brown, C., Gallo, C., Chiapa, A., Mauricio, A., & Jones, S. (2018). "Home practice is the program": Parents' practice of program skills as predictors of outcomes in the New Beginnings Program effectiveness trial. *Prevention Science, 19*(5), 663-673. https://doi.org/10.1007/s11121-016-0738-0

Betancourt, T., & Khan, K. (2008). The mental health of children affected by armed conflict: Protective processes and pathways to resilience. *International Review of Psychiatry, 20*(3), 317-328. https://doi.org/10.1080/09540260802090363

Bradley, R., & Corwyn, R. (2005). Caring for children around the world: A view from HOME. *International Journal of Behavioral Development, 29*(6), 468-478. https://doi.org/10.1080/01650250500146925

Breitenstein, S., Gross, D., & Christophersen, R. (2014). Digital delivery methods of parenting training interventions: A systematic review. *Worldviews on Evidence-Based Nursing, 11*(3), 168-176. https://doi.org/10.1111/wvn.12040

Bursac, Z., Gauss, C., Williams, D., & Hosmer, D. (2008). Purposeful selection of variables in logistic regression. *Source Code for Biology and Medicine, 3*(1), 17. https://doi.org/10.1186/1751-0473-3-17

Calam, R., Bolton, C., & Roberts, J. (2002). Maternal expressed emotion, attributions and depression and entry into therapy for children with behaviour problems. *The British Journal of Clinical Psychology, 41*(2). https://doi.org/10.1348/014466502163985

Cameron, A. C., & Miller, D. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources, 50*(2), 317-372. https://doi.org/10.3368/jhr.50.2.317

Chacko, A., Jensen, S., Lowry, L., Cornwell, M., Chimklis, A., Chan, E., Lee, D., Pulgarin, B., Jensen, S., & Lowry, L. (2016). Engagement in behavioral parent training: Review of the literature and implications for practice. *Clinical Child & Family Psychology Review, 19*(3), 204-215. https://doi.org/10.1007/s10567-016-0205-2

Chamberlain, P., Price, J., Reid, J., Landsverk, J., Fisher, P., & Stoolmiller, M. (2006). Who disrupts from placement in foster and kinship care? *Child Abuse & Neglect, 30*(4), 409-424. https://doi.org/10.1016/j.chiabu.2005.11.004

Chamberlain, P., & Reid, J. (1987). Parent observation and report of child symptoms. *Behavioral Assessment, 9*(1), 97-109. https://psycnet.apa.org/record/1988-12602-001

Chen, M., & Chan, K. (2016). Effects of parenting programs on child maltreatment prevention: A meta-analysis. *Trauma, Violence, & Abuse, 17*(1), 88-104. https://doi.org/10.1177/1524838014566718

Chiesa, A., Kallechey, L., Harlaar, N., Rashaan Ford, C., Garrido, E., Betts, W., & Maguire, S. (2018). Intimate partner violence victimization and parenting: A systematic review. *Child Abuse & Neglect, 80,* 285-300. https://doi.org/10.1016/j.chiabu.2018.03.028

Clerkin, S., Halperin, J., Marks, D., & Policaro, K. (2007). Psychometric properties of the Alabama Parenting Questionnaire – Preschool revision. *Journal of Clinical Child & Adolescent Psychology, 36*(1), 19-28. https://doi.org/10.1080/15374410709336565

Cluver, L., Meinck, F., Steinert, J., Shenderovich, Y., Doubt, J., Herrero Romero, R., Lombard, C. J., Redfern, A., Ward, C. L., Tsoanyane, S., Nzima, D., Sibanda, N., Wittesaele, C., De Stone, S., Boyes, M. E., Catanho, R., Lachman, J., Salah, N., Nocuza, M., & Gardner, F. (2018). Parenting for Lifelong Health: A pragmatic cluster

randomised controlled trial of a non-commercialised parenting programme for adolescents and their families in South Africa. *BMJ Global Health, 3*(1). https://doi.org/10.1136/bmjgh-2017-000539

Cluver, L., Orkin, M., Campeau, L., Toska, E., Webb, D., Carlqvist, A., & Sherr, L. (2019). Improving lives by accelerating progress towards the UN Sustainable Development Goals for adolescents living with HIV: A prospective cohort study. *The Lancet Child & Adolescent Health, 3*(4), 245-254. https://doi.org/https://doi.org/10.1016/S2352-4642(19)30033-1

Collins, L. (2018). *Optimization of behavioral, biobehavioral, and biomedical Interventions: The multiphase optimization strategy (MOST)*. Springer International Publishing.

Connell, A. (2009). Employing complier average causal effect analytic methods to examine effects of randomized encouragement trials. *The American Journal of Drug and Alcohol Abuse, 35*(4), 253-259. https://doi.org/10.1080/00952990903005882

Connolly, P. (2005). Children, assessments and computer-assisted interviewing. *Child Abuse Review, 14*(6), 407-414. https://doi.org/10.1002/car.922

Cortina, J. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98-104. https://doi.org/10.1037/0021-9010.78.1.98

Coxe, S., West, S., & Aiken, L. (2009). The analysis of count data: A gentle introduction to poisson regression and its alternatives. *Journal of Personality Assessment, 91*(2), 121-136. https://doi.org/10.1080/00223890802634175

Craig, P., Dieppe, P., Macintyre, S., Michie, S., Nazareth, I., Petticrew, M., & Medical Research Council. (2008). Developing and evaluating complex interventions: The new Medical Research Council guidance. *British Medical Journal, 337*, a1655. https://doi.org/10.1136/bmj.a1655

Dadds, M. R., Maujean, A., & Fraser, J. A. (2003). Parenting and conduct problems in children: Australian data and psychometric properties of the Alabama Parenting Questionnaire. *Australian Psychologist, 38*(3), 238-241. https://doi.org/10.1080/00050060310001707267

Dakof, G., Tejeda, M., & Liddle, H. (2001). Predictors of engagement in adolescent drug abuse treatment. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(3), 274-281. https://doi.org/10.1097/00004583-200103000-00006

Dane, A., & Schneider, B. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control. *Clinical Psychology Review, 18*(1), 23-45. https://doi.org/10.1016/S0272-7358(97)00043-3

Davies, M., & Morgan, A. (2005). Using computer-assisted self-interviewing (CASI) questionnaires to facilitate consultation and participation with vulnerable young people. *Child Abuse Review, 14*(6), 389-406. https://doi.org/10.1002/car.925

De Los Reyes, A., Augenstein, T. M., Wang, M., Thomas, S. A., Drabick, D. A. G., Burgers, D. E., & Rabinowitz, J. (2015). The validity of the multi-informant approach to assessing child and adolescent mental health. *Psychological Bulletin, 141*(4), 858-900. https://doi.org/10.1037/a0038498

DeGarmo, D. S., & Jones, J. A. (2019). Fathering Through Change (FTC) intervention for single fathers: Preventing coercive parenting and child problem behaviors. *Development and Psychopathology, 31*(5), 1801-1811. https://doi.org/10.1017/S0954579419001019

Del Vecchio, T., Jerusalmi, D., & Terjesen, M. (2017). Psychometric characteristics of the Parenting Scale in a Vietnamese sample. *International Journal of Psychology, 52*(6), 482-490. https://doi.org/10.1002/ijop.12242

Deng, L., & Chan, W. (2016). Testing the difference between reliability coefficients alpha and omega. *Educational and Psychological Measurement, 77*(2), 185-203. https://doi.org/0.1177/0013164416658325

Dishion, T. J., Brennan, L. M., Shaw, D. S., McEachern, A. D., Wilson, M. N., & Jo, B. (2014). Prevention of problem behavior through annual family check-ups in early childhood: intervention effects from home to early elementary school. *Journal of Abnormal Child Psychology, 42*(3), 343-354. https://doi.org/10.1007/s10802-013-9768-2

Dodge, K. A. (2001). The science of youth violence prevention: Progressing from developmental epidemiology to efficacy to effectiveness to public policy. *American Journal of Preventive Medicine, 20*(1), 63-70. https://doi.org/10.1016/S0749-3797(00)00275-0

Doubt, J., Loening-Voysey, H., Blanc, D., Cluver, L., Byrne, J., Petersen, T., & Unicef. (2018). *Delivering a parenting programme in rural South Africa: The local child and youth care worker experience*. https://www.unicef-irc.org/publications/pdf/IWP2018-01%20-%20Parenting%20Programme%20in%20Rural%20South%20Africa.pdf

Dumas, J. E., Nissley-Tsiopinis, J., & Moreland, A. D. (2007). From intent to enrollment, attendance, and participation in preventive parenting groups. *Journal of Child and Family Studies, 16*(1), 1-26. https://doi.org/10.1007/s10826-006-9042-0

Dunn, G., Maracy, M., & Tomenson, B. (2005). Estimating treatment effects from randomized clinical trials with noncompliance and loss to follow-up: The role of instrumental variable methods. *Statistical Methods in Medical Research, 14*(4), 369-395. https://doi.org/10.1191/0962280205sm403oa

Dunn, T., Baguley, T., & Brunsden, V. (2014). From alpha to omega: A practical solution to the pervasive problem of internal consistency estimation *British Journal of Psychology, 105*(3), 399-412. https://doi.org/10.1111/bjop.12046

Dunne, M. P., Zolotor, A. J., Runyan, D. K., Andreva-Miller, I., Choo, W. Y., Dunne, S. K., Gerbaka, B., Isaeva, O., Jain, D., Kasim, M. S., Macfarlane, B., Mamyrova, N., Ramirez, C., Volkova, E., & Youssef, R. (2009). ISPCAN Child Abuse Screening Tools Retrospective version (ICAST-R): Delphi study and field testing in seven countries. *Child Abuse & Neglect, 33*(11), 815-825. https://doi.org/10.1016/j.chiabu.2009.09.005

Duppong-Hurley, K., Hoffman, S., Barnes, B., & Oats, R. (2016). Perspectives on engagement barriers and alternative delivery formats from non-completers of a community- run parenting program. *Journal of Child & Family Studies, 25*(2), 545-552. https://doi.org/10.1007/s10826-015-0253-0

Durlak, J. A., & DuPre, E. P. (2008). Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American Journal of Community Psychology, 41*(3/4), 327-350. https://doi.org/10.1007/s10464-008-9165-0

Elgar, F., Waschbusch, D., Dadds, M., & Sigvaldason, N. (2007). Development and validation of a short form of the Alabama Parenting Questionnaire [Article]. *Journal of Child & Family Studies, 16*(2), 243-259. https://doi.org/10.1007/s10826-006-9082-5

Elliott, D. S., & Mihalic, S. (2004). Issues in disseminating and replicating effective prevention programs. *Prevention Science, 5*(1), 47-53. https://doi.org/10.1023/b:Prev.0000013981.28071.52

Errázuriz, P., Cerfogli, C., Moreno, G., & Soto, G. (2016). Perception of Chilean parents on the Triple P program for improving parenting practices. *Journal of Child and Family Studies, 25*(11), 3440-3449. https://doi.org/10.1007/s10826-016-0492-8

Essau, C. A., Sasagawa, S., & Frick, P. J. (2006). Psychometric properties of the Alabama Parenting Questionnaire. *Journal of Child & Family Studies, 15*(5), 595-614. https://doi.org/10.1007/s10826-006-9036-y

Eyberg, S. M., & Ross, A. W. (1978). Assessment of child behavior problems: The validation of a new inventory. *Journal of Clinical Child Psychology, 7*(2), 113-116. https://doi.org/10.1080/15374417809532835

Fang, X., Brown, D. S., Florence, C. S., & Mercy, J. A. (2012). The economic burden of child maltreatment in the United States and implications for prevention. *Child Abuse & Neglect, 36*(2), 156-165. https://doi.org/10.1016/j.chiabu.2011.10.006

Fang, X., Fry, D., Brown, D., Mercy, J., Dunne, M., Butchart, A., Corso, P., Maynzyuk, K., Dzhygyr, Y., Chen, Y., McCoy, A., & Swales, D. (2015). The burden of child maltreatment in the East Asia and Pacific region. *Child Abuse & Neglect, 42*, 146-162. https://doi.org/10.1016/j.chiabu.2015.02.012

Fang, X., Zheng, X., Fry, D., Ganz, G., Casey, T., Hsiao, C., & Ward, C. L. (2017). The economic burden of violence agaisnt children in South Africa. *Internaltional Journal of Environmental Research and Public Health, 14*(11), 1431. https://doi.org/10.3390/ijerph14111431

Featherstone, B., Morris, K., White, S., & White, S. (2014). *Re-imagining child protection: Towards humane social work with families*. Policy Press.

Fixsen, D., Naoom, S., Blase, K., & Friedman, R. (2005). *Implementation research: A synthesis of the literature.*

http://citeseerx.ist.psu.edu/viewdoc/download;jsessionid=576877CD03619E82CCE728755B0AC0B5?doi=10.1.1.610.6226&rep=rep1&type=pdf

Flay, B., Biglan, A., Boruch, R., Castro, F., Gottfredson, D., Kellam, S., Mościcki, E., Schinke, S., Valentine, J., & Ji, P. (2005). Standards of evidence: Criteria for efficacy, effectiveness and dissemination. *Prevention Science, 6*(3), 151-175. https://doi.org/10.1007/s11121-005-5553-y

Frangakis, C., & Rubin, D. (1999). Addressing complications of intention-to-treat analysis in the combined presence of all-or-none treatment-noncompliance and subsequent missing outcomes. *Biometrika, 86*(2), 365-379. https://doi.org/10.1093/biomet/86.2.365

Frantz, I., Foran, H., Lachman, J., Jansen, E., Hutchings, J., Băban, A., Fang, X., Gardner, F., Lesco, G., Raleva, M., & Ward, C. (2019). Prevention of child mental health problems in Southeastern Europe: A multicentre sequential study to adapt, optimise and test the parenting programme 'Parenting for Lifelong Health for Young Children', protocol for stage 1, the feasibility study. *BMJ Open, 9*(1). https://doi.org/10.1136/bmjopen-2018-026684

Fry, D., Fang, X., Elliott, S., Casey, T., Zheng, X., Li, J., Florian, L., & 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. https://doi.org/10.1016/j.chiabu.2017.06.021

Furlong, M., McGilloway, S., Bywater, T., Hutchings, J., Smith, S., & Donnelly, M. (2013). Cochrane Review: Behavioural and cognitive-behavioural group-based parenting programmes for early-onset conduct problems in children aged 3 to 12 years (Review). *Evidence-Based Child Health: A Cochrane Review Journal, 8*(2), 318-692. https://doi.org/10.1002/ebch.1905

Garcia-Moreno, C., Jansen, H., Ellsberg, M., Heise, L., & Watts, C. (2006). Prevalence of intimate partner violence: Findings from the WHO multi-country study on women's health and domestic violence. *The Lancet, 368*(9543), 1260-1269. https://doi.org/10.1016/S0140-6736(06)69523-8

Gardner, F., Burton, J., & Klimes, I. (2006). Randomised controlled trial of a parenting intervention in the voluntary sector for reducing child conduct problems: Outcomes and mechanisms of change. *Journal of Child Psychology and Psychiatry, 47*(11), 1123-1132. https://doi.org/10.1111/j.1469-7610.2006.01668.x

Gardner, F., McCoy, A., Lachman, J., Melendez-Torres, G., Ward, C., Cheah, P., & Topanya, S. (forthcoming). Randomized trial of a parenting intervention embedded within the public health system to reduce violence against children in Thailand.

Gardner, F., Shaw, D., Dishion, T., Burton, J., & Supplee, L. (2007). Randomized prevention trial for early conduct problems: Effects on proactive parenting and links to toddler disruptive behavior. *Journal of Family Psychology, 21*(3), 398-406. https://doi.org/10.1037/0893-3200.21.3.398

Gorman-Smith, D., Tolan, P., Henry, D., Leventhal, A., Schoeny, M., Lutovsky, K., & Quintana, E. (2002). Predictors of participation in a family-focused preventive intervention for substance use. *Psychology of Addictive Behaviors, 16*(4), S55-S64. https://doi.org/10.1037/0893-164X.16.4S.S55

Gross, D., Garvey, C., Julion, W., Fogg, L., Tucker, S., & Mokros, H. (2009). Efficacy of the Chicago parent program with low-income African American and latino parents of young children. *Prevention Science, 10*(1), 54-65. https://doi.org/10.1007/s11121-008-0116-7

Guo, M., Morawska, A., & Sanders, M. (2016). A randomized controlled trial of group Triple P with Chinese parents in Mainland China. *Behavior Modification, 40*(6), 825-851. https://doi.org/10.1177/0145445516644221

Hall, C. M., & Bierman, K. (2015). Technology-assisted interventions for parents of young children: Emerging practices, current research, and future directions. *Early Childhood Research Quarterly, 33*, 21-32. https://doi.org/10.1016/j.ecresq.2015.05.003

Head, E. (2009). The ethics and implications of paying participants in qualitative research. *International Journal of Social Research Methodology, 12*(4), 335-344. https://doi.org/10.1080/13645570802246724

Heinrichs, N., Bertram, H., Kuschel, A., & Hahlweg, K. (2005). Parent recruitment and retention in a universal prevention program for child behavior and emotional problems: Barriers to research and program participation. *Prevention Science, 6*(4), 275-286. https://doi.org/10.1007/s11121-005-0006-1

Hillis, S., Mercy, J., Amobi, A., & Kress, H. (2016). Global prevalence of past-year violence against children: A systematic review and minimum estimates. *Pediatrics, 137*(3). https://doi.org/10.1542/peds.2015-4079

Hogan, T., Benjamin, A., & Brezinski, K. (2000). Reliability methods: A note on the frequency of use of various types. *Educational and Psychological Measurement, 60*(4), 523-531. https://doi.org/10.1177/00131640021970691

Hosmer, D., & Lemeshow, S. (2000). *Applied logistic regression*. John Wiley & Sons.

Huang, S., Cordova, D., Estrada, Y., Brincks, A., Asfour, L., & Prado, G. (2014). An application of the complier average causal effect analysis to examine the effects of a family intervention in reducing illicit drug use among high-risk Hispanic adolescents. *Family Process, 53*(2), 336-347. https://doi.org/10.1111/famp.12068

Hughes, K., Bellis, M., Hardcastle, K., Sethi, D., Butchart, A., Mikton, C., Jones, L., & Dunne, M. (2017). The effect of multiple adverse childhood experiences on health: A systematic review and meta-analysis. *The Lancet Public Health, 2*(8), e356-e366. https://doi.org/10.1016/S2468-2667(17)30118-4

Hutchings, J., Bywater, T., & Daley, D. (2007). Early prevention of conduct disorder: How and why did the North and Mid Wales Sure Start study work? *Journal of Children's Services, 2*(2), 4-14. https://doi.org/10.1108/17466660200700012

Hutchings, J., Bywater, T., Daley, D., Gardner, F., Whitaker, C., Jones, K., Eames, C., & Edwards, R. T. (2007). Parenting intervention in Sure Start services for children at risk of developing conduct disorder: pragmatic randomised controlled trial. *BMJ, 334*(7595), 678. https://doi.org/10.1136/bmj.39126.620799.55

Inter-Agency Standing Committee (IASC). (2007). *IASC guildines on mental health and psychosocial support in emergency settings*. https://www.who.int/mental_health/emergencies/9781424334445/en/

Isaranurug, S., Nitirat, P., Chauytong, P., & Wongarsa, C. (2001). Factors relating to the aggressive behavior of primary caregiver toward a child. *Journal of the Medical Association of Thailand, 84*(10), 1481-1489. http://europepmc.org/abstract/MED/11804260

Janowski, R., Wessels, I., Bojo, S., Monday, F., Maloney, K., Achut, V., Oliver, D., Lachman, J., Cluver, L., & Ward, C. (2020). Transferability of evidence-based parenting programs to routine implementation in postconflict South Sudan. *Research on Social Work Practice, 20*(10), 1-12. https://doi.org/10.1177/1049731520932986

Jeong, J., Adhia, A., Bhatia, A., McCoy, D., & Yousafzai, A. (2020). Intimate partner violence, maternal and paternal parenting, and early child development. *Pediatrics, 145*(6), e20192955. https://doi.org/10.1542/peds.2019-2955

Jo, B. (2002). Statistical power in randomized intervention studies with noncompliance. *Psychological Methods, 7*(2), 178-193. https://doi.org/10.1037/1082-989X.7.2.178

Jo, B., Asparouhov, T., Muthén, B., Ialongo, N., & Brown, C. (2008). Cluster randomized trials with treatment noncompliance. *Psychological Methods, 13*(1), 1-18. https://doi.org/10.1037/1082-989X.13.1.1

Jo, B., & Muthén, B. (2001). Modeling of intervention effects with noncompliance: A latent variable approach for randomized trials. In *New developments and techniques in Structural Equation Modeling* (pp. 57-87). Mahwah, NJ: Erlbaum, Inc.

Jordans, M., Pigott, H., & Tol, W. (2016). Interventions for children affected by armed conflict: A systematic review of mental health and psychosocial support in low- and middle-income countries. *Current Psychiatry Reports, 18*(1), 9. https://doi.org/10.1007/s11920-015-0648-z

Kazdin, A., Mazurick, J., & Bass, D. (1993). Risk for attrition in treatment of antisocial children and families. *Journal of Clinical Child Psychology 22*(1), 2-16. https://doi.org/10.1207/s15374424jccp2201_1

Kazdin, A. E. (2000). Perceived barriers to treatment participation and treatment acceptability among antisocial children and their families. *Journal of Child & Family Studies*(9), 157-174. https://doi.org/10.1023/a:1009414904228

Kim, C., Choi, H., Ko, H., & Park, C. (2020). Agreement between parent proxy reports and self-reports of adolescent emotional distress. *The Journal of School Nursing, 36*(2), 104-111. https://doi.org/10.1177/1059840518792073

Kim, J., Schulz, W., Zimmermann, T., & Hahlweg, K. (2018). Parent–child interactions and child outcomes: Evidence from randomized intervention. *Labour Economics, 54*, 152-171. https://doi.org/10.1016/j.labeco.2018.08.003

Kleiber, C., & Zeileis, A. (2008). *Applied Econometrics with R*. Springer-Verlag. http://CRAN.R-project.org/package=AER

Knerr, W., Gardner, F., & Cluver, L. (2013). Improving positive parenting skills and reducing harsh and abusive parenting in low- and middle-income countries: A systematic review. *Prevention Science, 14*(4), 352-363. https://doi.org/10.1007/s11121-012-0314-1

Knox, C., Lall, R., Hansen, Z., & Lamb, S. (2014). Treatment compliance and effectiveness of a cognitive behavioural intervention for low back pain: A complier average causal effect approach to the BeST data set. *BMC Musculoskeletal Disorders, 15*(1), 17. https://doi.org/10.1186/1471-2474-15-17

Knox, M. S., Burkhart, K., & Hunter, K. E. (2011). ACT against violence parents raising safe kids program: Effects on maltreatment-related parenting behaviors and beliefs. *Journal of Family Issues, 32*(1), 55-74. https://doi.org/10.1177/0192513x10370112

Kogan, S., Lei, M., Brody, G., Futris, T., Sperr, M., & Anderson, T. (2016). Implementing family-centered prevention in rural African American communities: A randomized effectiveness trial of the Strong African American Families Program. *Prevention science, 17*(2), 248-258. https://doi.org/10.1007/s11121-015-0614-3

Labadarios, D., Steyn, N., Maunder, E., MacIntryre, U., Gericke, G., Swart, R., Huskisson, J., Dannhauser, A., Vorster, H., Nesmvuni, A., & Nel, J. (2005, Aug). The National Food Consumption Survey (NFCS): South Africa, 1999. *Public Health Nutrition, 8*(5), 533-543. https://doi.org/10.1079/phn2005816

Lachman, J., Alampay, L., Alinea, M., Madrid, B., Ward, C., Hutchings, J., Landoy, B., Garilao, M., & Gardner, F. (forthcoming). Integrating parent management training within a conditional cash transfer system for low-income families: Results from a small-scale randomized controlled trial in the Philippines.

Lachman, J., Cluver, L., Ward, C. L., Hutchings, J., Mlotshwa, S., Wessels, I., & Gardner, F. (2017). Randomized controlled trial of a parenting program to reduce the risk of child maltreatment in South Africa. *Child Abuse & Neglect*, *72*, 338-351. https://doi.org/10.1016/j.chiabu.2017.08.014

Lachman, J., Sherr, L., Cluver, L., Ward, C., Hutchings, J., & Gardner, F. (2016). Integrating evidence and context to develop a parenting program for low-income families in South Africa. *Journal of Child and Family Studies*, *25*(7), 2337-2352. https://doi.org/10.1007/s10826-016-0389-6

Lachman, J. M., Heinrichs, N., Jansen, E., Brühl, A., Taut, D., Fang, X., Gardner, F., Hutchings, J., Ward, C., Williams, M. E., Raleva, M., Båban, A., Lesco, G., & Foran, H. (2019). Preventing child mental health problems through parenting interventions in Southeastern Europe (RISE): Protocol for a multi-country cluster randomized factorial study. *Contemporary Clinical Trials*, *86*, 105855. https://doi.org/10.1016/j.cct.2019.105855

Lansford, J., Sharma, C., Malone, P., Woodlief, D., Dodge, K., Oburu, P., Pastorelli, C., Skinner, A., Sorbring, E., Tapanya, S., Tirado, L. M., Zelli, A., Al-Hassan, S., Alampay, L., Bacchini, D., Bombi, A., Bornstein, M., Chang, L., Deater-Deckard, K., & Di Giunta, L. (2014). Corporal punishment, maternal warmth, and child adjustment: A longitudinal study in eight countries. *Journal of Clinical Child and Adolescent Psychology 43*(4), 670-685. https://doi.org/10.1080/15374416.2014.893518

Laxman, D., Higginbotham, B., & Bradford, K. (2019). Predictors of attrition and attendance in a fatherhood education program. *Children and Youth Services Review*, *103*, 287-297. https://doi.org/10.1016/j.childyouth.2019.05.007

Leijten, P., Gardner, F., Melendez-Torres, G., van Aar, J., Hutchings, J., Schulz, S., Knerr, W., & Overbeek, G. (2018). Meta-analyses: Key parenting program components for disruptive child behavior. *Journal of the American Academy of Child & Adolescent Psychiatry, 58*(2), 180-190. https://doi.org/10.1016/j.jaac.2018.07.900

Little, M., Berry, V., Morpeth, L., Blower, S., Axford, N., Taylor, R., Bywater, T., Lehtonen, M., & Tobin, K. (2012). The impact of three evidence-based programmes delivered in public systems in Birmingham, UK. *International Journal of Conflict and Violence, 6*(2), 260-272. http://clok.uclan.ac.uk/6824/1/6824_Berry.pdf

Little, R. J., & Yau, L. (1998). Statistical techniques for analyzing data from prevention trials: Treatment of no-shows using Rubin's causal model. *Psychological Methods, 3*(2), 147-159. https://doi.org/10.1037/1082-989X.3.2.147

Lovibond, S., & Lovibond, P. (2004). *Manual for the depression anxiety stress scales* (Fourth ed.). Psychology Foundation of Australia.

Lundahl, B., Tollefson, D., Risser, H., & Lovejoy, M. C. (2008). A meta-analysis of father involvement in parent training. *Research on Social Work Practice, 18*(2), 97-106. https://doi.org/10.1177/1049731507309828

Mackenzie, M., Booij, A., Hutchings, J., & Lachman, J. (2019). *Parenting for Lifelong Health for Young Children: Facilitator Assessment Tool Manual*.

Madans, J., Loeb, M., & Altman, B. (2011). Measuring disability and monitoring the UN convention on the rights of persons with disabilities: The work of the Washington Group on disability statistics. *BMC Public Health, 11*(4), S4. https://doi.org/10.1186/1471-2458-11-S4-S4

McCoy, A., Melendez-Torres, G. J., & Gardner, F. (2020). Parenting interventions to prevent violence against children in low- and middle-income countries in East and Southeast

Asia: A systematic review and multi-level meta-analysis. *Child Abuse & Neglect, 103*. https://doi.org/10.1016/j.chiabu.2020.104444

McEachern, A., Dishion, T., Weaver, C., Shaw, D., Wilson, M., & Gardner, F. (2012). Parenting Young Children (PARYC): Validation of a self-report parenting measure. *Journal of Child & Family Studies, 21*(3), 498-511. https://doi.org/10.1007/s10826-011-9503-y

McNamee, R. (2009). Intention to treat, per protocol, as treated and instrumental variable estimators given non-compliance and effect heterogeneity. *Statistics in Medicine, 28*(21), 2639-2652. https://doi.org/10.1002/sim.3636

McNeish, D. (2018). Thanks coefficient alpha, we'll take it from here. *Psychological Methods, 23*(3), 412-433. https://doi.org/10.1037/met0000144

Meinck, F., Boyes, M. E., Cluver, L., Ward, C. L., Schmidt, P., DeStone, S., & Dunne, M. (2018). Adaptation and psychometric properties of the ISPCAN Child Abuse Screening Tool for use in trials (ICAST-Trial) among South African adolescents and their primary caregivers. *Child Abuse & Neglect, 82*, 45-58. https://doi.org/10.1016/j.chiabu.2018.05.022

Mejia, A., Calam, R., & Sanders, M. (2015). A pilot randomized controlled trail of a brief parenting intervention in low-resource settings in Panama. *Prevention Science, 16*(5), 707-717. https://doi.org/10.1007/s11121-015-0551-1

Menting, A., Orobio de Castro, B., & Matthys, W. (2013). Effectiveness of The Incredible Years parent training to modify disruptive and prosocial child behavior: A metaanalytic review. *Clinical Psychology Review, 33*, 901-913. https://doi.org/10.1016/j.cpr.2013.07.006

Mickey, R., & Greenland, S. (1989). The impact of confounder selection criteria on effect estimation. *American Journal of Epidemiology, 129*(1), 125-137. https://doi.org/10.1093/oxfordjournals.aje.a115101

Mikton, C. (2012). Two challenges to importing evidence-based child maltreatment prevention programs developed in high-income countries to low-and middle-income countries: Generalizability and affordability. In H. Dubowits (Ed.), World Perspectives on Child Abuse and Neglect International Society for the Prevention of Child Abuse and Neglect. http://www.ispcan.org/page/WP_12

Mikton, C., & Buchart, A. (2009). Child maltreatment prevention: A systematic review of reviews. *Bulletin of the World Health Organization, 87*, 353-361. https://doi.org/10.2471/BLT.08.057075

Montgomery, M., Gragnolati, M., Burke, K., & Paredes, E. (2000). Measuring living standards with proxy variables. *Demography, 37*(2), 155-174. https://doi.org/10.2307/2648118

Moore, M. (2020). *Internet penetration in Southeast Asia 2020, by country*. https://www.statista.com/statistics/487965/internet-penetration-in-southeast-asian-countries/

Murphy, K, Rodrigues, K., Costigan, J., & Annan, J. (2017). Raising children in conflict: An integrative model of parenting in war. *Peace and Conflict: Journal of Peace Psychology, 23*(1), 46-57. https://doi.org/10.1037/pac0000195

Mytton, J., Ingram, J., Manns, S., & Thomas, J. (2014). Facilitators and barriers to engagement in parenting programs: A qualitative systematic review. *Health Education & Behavior, 41*(2), 127-137. https://doi.org/10.1177/1090198113485755

National Statistical Office & United Nations Children's Fund. (2016). *Thailand Multiple Indicator Cluster Survey 2015-2016*. http://www.unicef.org/thailand/sites/unicef.org.thailand/files/2018-06/Thailand_MICS_Full_Report_EN_0.pdf

National Statistical Office and Ministry of Information and Communication Technology. (2016). *The 2015 Household Socioeconomic Survey: Whole Kingdom*. http://www.nso.go.th/sites/2014en/Pages/survey/Social/Household/The%20Household.aspx

National Statistical Office Thailand. (2015). *The 2015 household socio-economic survey Northeastern region.* http://web.nso.go.th/en/survey/house_seco/data/Full_Report2015_ne.pdf

Nix, R., Bierman, K., & McMohon, R. (2009). How attendance and quality of participation affect treatment response to parent management training. *Jounrnal of Consulting and Clinical Psychology*, *77*(3), 429-438. https://doi.org/10.1037/a0015028

Ogden, T., & Fixsen, D. (2014). Implementation science: A brief overview and a look ahead. *Zeitschrift für Psychologie*, *222*(1), 4-11. https://doi.org/10.1027/2151-2604/a000160

Panter-Brick, C., Burgess, A., Eggerman, M., McAllister, F., Pruett, K., & Leckman, J. (2014). Practitioner review: Engaging fathers – recommendations for a game change in parenting interventions based on a systematic review of the global evidence. *Journal of Child Psychology and Psychiatry*, *55*(11), 1187-1212. https://doi.org/10.1111/jcpp.12280

Patterson, G., Chamberlain, P., & Reid, J. (1982). A comparative evaluation of a parent-training program. *Behavior Therapy*, *13*(5), 638-650. https://doi.org/10.1016/S0005-7894(82)80021-X

Patterson, J., Mockford, C., Barlow, J., Pyper, C., & Stewart-Brown, S. (2002). Need and demand for parenting programmes in general practice. *Archives of Disease in Childhood, 87*(6), 468-471. https://doi.org/10.1136/adc.87.6.468

Peters, S., Calam, R., & Harrington, R. (2005). Maternal attributions and expressed emotion as predictors of attendance at parent management training. *Jounral of Child Psychology and Psychiatry and Allied Disciplines*(46), 436-448. https://doi.org/10.1111/j.1469-7610.2004.00365.x

Philippine Statistics Authority. (2015). *2015 census of population.* . http://www.psa.gov.ph/statistics/census/2015-census-of-population

Philippine Statistics Authority. (2018). *The 2018 Family Income and Expenditure Survey*. http://www.psa.gov.ph/sites/default/files/FIES%202018%20Final%20Report.pdf

Phillips, A., Gomez, G., Boily, M., & Garnett, G. (2010). A systematic review and meta-analysis of quantitative interviewing tools to investigate self-reported HIV and STI associated behaviours in low- and middle-income countries. *International Journal of Epidemiology, 39*(6), 1541-1555. https://doi.org/10.1093/ije/dyq114

Piquero, A., Farrington, D., Welsh, B., Tremblay, R., & Jennings, W. (2009). Effects of early family/parent training programs on antisocial behavior and delinquency. *Journal of Experimental Criminology, 5*(2), 83-120. https://doi.org/10.1007/s11292-009-9072-x

Provincial Community Development Office of Udon Thani. (2017). *Presentation on the Results of Developing Population Quality of Life: Udon Thani* http://udon.cdd.go.th/services/ebmn-cddudon-report-60

R Core Team. (2018). *R: A language and enironment for statistical computing* In (Version 3.6.1) R Foundation for Statistical Computing. https://www.R-project.org/

Reitman, D., Currier, R. O., Hupp, S., Rhode, P., Murphy, M., & O'Callaghan, P. (2001). Psychometric characteristics of the Parenting Scale in a head start population. *Journal*

of Clinical Child Psychology, 30(4), 514-524. https://doi.org/10.1207/s15374424jccp3004_08

Revelle, W. (2018). *Procedures for Personality and Psychological Research*. https://cran.r-project.org/web/packages=psych.

Revelle, W., & Zinbarg, R. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika, 74*(1), 145-154. https://doi.org/10.1007/s11336-008-9102-z

Reyno, S., & McGrath, P. (2006). Predictors of parent training efficacy for child externalizing behavior problems – A meta-analytic review. *Journal for Child Psychology and Psychiatry, 47*(1), 99-111. https://doi.org/10.1111/j.1469-7610.2005.01544.x

Runyan, D., Shankar, V., Hassan, F., Hunter, W., Jain, D., Paula, C. S., Bangdiwala, S., Ramiro, L., Muñoz, S., Vizcarra, B., & Bordin, I. (2010). International variations in harsh child discipline. *Pediatrics, 126*(3), e701-e711. https://doi.org/10.1542/peds.2008-2374

Salari, R., & Filus, A. (2017). Using the Health Belief Model to explain mothers' and fathers' intention to participate in universal parenting programs. *Prevention Science, 18*(1), 83-94. https://doi.org/10.1007/s11121-016-0696-6

Sanders, M., Stallman, H., & McHale, M. (2011). Workplace Triple P: A controlled evaluation of a parenting intervention for working parents. *Journal of Family Psychology, 25*(4), 581. https://doi.org/10.1037/a0024148

Schafer, M. (2020). *The effect of parenting interventions on intimate partner violence: A systematic review and meta-analysis* (inpublished master's thesis). University of Oxford. Oxford, United Kingdom.

Schochet, P., & Chiang, H. (2009). *Technical methods report: Estimation and identification of the complier average causal effect parameter in education RCTs. NCEE 2009-4040*. National Center for Education Evaluation and Regional Assistance. https://files.eric.ed.gov/fulltext/ED504795.pdf

Schochet, P., & Chiang, H. (2011). Estimation and identification of the complier average causal effect parameter in education RCTs. *Journal of Educational and Behavioral Statistics*, *36*(3), 307-345. https://doi.org/10.3102/1076998610375837

Schochet, P., Puma, M., & Deke, J. (2014). *Understanding variation in treatment effects in education impact evaluations: An overview of quantitative methods*. I. o. E. S. Department of Education, National Center for Education Evaluation and Regional Assistance, Analytic Technical Assistance and Development. http://ies.ed.gov/ncee/edlabs

Schoenfelder, E., Sandler, I., Millsap, R., Wolchik, S., Berkel, C., & Ayers, T. (2013). Caregiver responsiveness to the family bereavement program: What predicts responsiveness? What does responsiveness predict? *Prevention Science*, *14*(6), 545-556. https://doi.org/10.1007/s11121-012-0337-7

Schulz, K., & Grimes, D. (2002). Sample size slippages in randomised trials: exclusions and the lost and wayward. *The Lancet, 359*(9308), 781-785. https://doi.org/10.1016/S0140-6736(02)07882-0

Sedwick, P. (2015). Intention to treat analysis versus per protocol analysis of trial data. *British Medical Journal, 350,* h681. https://doi.org/10.1136/bmj.h681

Shenderovich, Y., Eisner, M., Cluver, L., Doubt, J., Berezin, M., Majokweni, S., & Murray, A. (2019, April 01). Delivering a parenting program in South Africa: The impact of implementation on outcomes. *Journal of Child and Family Studies, 28*(4), 1005-1017. https://doi.org/10.1007/s10826-018-01319-y

Shenderovich, Y., Eisner, M., Cluver, L., Doubt, J., Berezin, M., Majokweni, S., & Murray, A. L. (2018). What affects attendance and engagement in a parenting program in South Africa? *Prevention Science, 19*, 977-986. https://doi.org/0.1007/s11121-018-0941-2

Shenderovich, Y., Lachman, J., Wessels, I., Ward, C., Gardner, F., Tomlinson, M., Oliver, D., Janowski, R., Martin, M., Okop, K., Sacolo-Gwebu, H., Ngcobo, L., Fang, Z., Alampay, L., Baban, A., Baumann, A., Benevides de Barros, R., Bojo, S., Buchart, A., Dippenaar, W., Exavery, A., Fang, X., Ferdinandi, I., Foran, H., Heinrichs, N., Hutchings, J., Kisyombe, D., Massetti, G., Mazak, J., Mbuyi, H., Singh, P., Polsky, K., Rakotomalala, S., Releva, M., Savo, R., & Cluver, L. (in submission). The science of scale for violence prevention: A new agenda for family strengthening in low- and middle-income countries. *Frontiers in Public Health*.

Sierau, S., Brand, T., Manly, J., Schlesier-Michel, A., Klein, A., Andreas, A., Garzón, L., Keil, J., Binser, M., von Klitzing, K., & White, L. (2017). A multisource approach to assessing child maltreatment from records, caregivers, and children. *Child Maltreatment*, *22*(1), 45-57. https://doi.org/10.1177/1077559516675724

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika, 74*(1), 107-120. https://doi.org/10.1007/s11336-008-9101-0

Sirles, E. (1990). Dropout from intake, diagnostics, and treatment. *Community Mental Health Journal*, *26*(4), 345-360. https://doi.org/10.1007/BF00752725

Sjolander, A., & Martinussen, T. (2019). Instrumental variable estimation with the R package ivtools. *Epidemiologic Methods*, *8*(1), 20180024. https://doi.org/10.1515/em-2018-0024

Smith, J., Berkel, C., Hails, K., Dishion, T., Shaw, D., & Wilson, M. (2018). Predictors of participation in the Family Check-Up Program: A randomized trial of yearly services

from age 2 to 10 years. *Prevention science, 19*(5), 652-662. https://doi.org/10.1007/s11121-016-0679-7

Spoth, R., Kavanagh, K., & Dishion, T. (2002). Family-centered preventive intervention science: Toward benefits to larger populations of children, yuth, and families. *Prevention Science, 3*, 145-153. https://doi.org/10.1023/A:1019924615322

Stanger, C., Ryan, S., Fu, H., & Budney, A. (2011). Parent training plus contingency management for substance abusing families: A complier average causal effects (CACE) analysis. *Drug and Alcohol Dependence, 118*(2-3), 119-126. https://doi.org/10.1016/j.drugalcdep.2011.03.007

Straus, M., Hamby, S., Boney-McCoy, S., & Sugarman, D. (1996). The revised Conflict Tactics Scales (CTS2): Development and preliminary psychometric data. *Journal of Family Issues, 17*(3), 283-316. https://doi.org/10.1177/019251396017003001

Stuart, E., Perry, D., Le, H., & Ialongo, N. (2008). Estimating intervention effects of prevention programs: Accounting for noncompliance. *Prevention Science 9*(4), 288-298. https://doi.org/10.1007/s11121-008-0104-y

The World Bank. (2020). *Thailand*. https://www.worldbank.org/en/country/thailand

Tol, W., Song, S., & Jordans, M. (2013). Annual research review: Resilience and mental health in children and adolescents living in areas of armed conflict - a systematic review of findings in low- an middle-income countries. *Journal for Child Psychology and Psychiatry, 54*(4), 445-460. https://doi.org/10.1111/jcpp.12053

Tran, T., Tran, T., & Fisher, J. (2013). Validation of the depression anxiety stress scales (DASS) 21 as a screening instrument for depression and anxiety in a rural community-based cohort of northern Vietnamese women. *BMC Psychiatry, 13*, 24-24. https://doi.org/10.1186/1471-244X-13-24

UNICEF. (2005). *Multiple indicator cluster survey manual, 2005: Monitoring the situation of children and women*. https://mics.unicef.org/files?job=W1siZiIsIjIwMTUvMDQvMDIvMDYvMzcvMDYvMTE5L011bHRpcGxlX0luZGljYXRvcl9DbHVzdGVyX1N1cnZleV9NYW51YWxfMjAwNS5wZGYiXV0&sha=dd2e54d1ddd61cdb

UNICEF. (2016). *National baseline study on violence against children: Philippines*. https://www.unicef.org/philippines/sites/unicef.org.philippines/files/2019-02/phl-nbsvac-resultssummary.pdf

UNICEF East Asia and Pacific Regional Office. (2014). *Violence against children in East Asia and the Pacific: A regional review and synthesis of findings*. Strengthening Child Protection Series. https://www.unicef.org/eap/reports/violence-against-children-east-asia-and-pacific

United Nations Development Programme. (2010). Human Security Today and Tomorrow: Thailand Human Development Report 2009. http://www.th.undp.org/content/thailand/en/home/library/human_development/2009NHDR.html

United Nations General Assembly. (2006). *Rights of the child: Note by the Secretary-General*. https://www.refworld.org/docid/453780fe0.html

van de Mortel, T. (2008). Faking it: social desirability response bias in self-report research. *Australian Journal of Advanced Nurcing*, *25*(4), 40. https://epubs.scu.edu.au/cgi/viewcontent.cgi?article=1001&context=hahs_pubs

Walker, S., Baker-Henningham, H., Chang, S., Powell, C., Lopez-Boo, F., & Grantham-Mcgregor, S. (2018). Implementation of parenting interventions through health services in Jamaica. *Vulnerable Children and Youth Studies, 13*(2), 127-141. https://doi.org/10.1080/17450128.2017.1395100

Ward, C., Gould, C., Kelly, J., & Mauff, K. (2015). Spare the rod and save the child: Assessing the impact of parenting on child behaviour and mental health. *South African Crime Quarterly, 51*, 9. https://doi.org/10.4314/sacq.v51i0.2

Ward, C., Wessels, I., Lachman, J., Hutchings, J., Cluver, L., Kassanjee, R., Nhapi, R., Little, F., & Gardner, F. (2020). Parenting for Lifelong Health for Young Children: A randomized controlled trial of a parenting program in South Africa to prevent harsh parenting and child conduct problems. *Journal of Child Psychology and Psychiatry, 61*(4), 503-512. https://doi.org/10.1111/jcpp.13129

Ware, J., Kosinski, M., & Keller, S. (1996). A 12-item short-form health survey: Construction of scales and preliminary tests of reliability and validity. *Medical Care, 34*(3), 220-233. www.jstor.org/stable/3766749

Weeland, J., Chhangur, R., van der Giessen, D., Matthys, W., de Castro, B., & Overbeek, G. (2017). Intervention effectiveness of The Incredible Years: New insights into sociodemographic and intervention-based moderators. *Behavior Therapy, 48*(1), 1-18. https://doi.org/10.1016/j.beth.2016.08.002

Weissinger, E., Caldwell, L., & Bandalos, D. (1992). Relation between intrinsic motivation and boredom in leisure time. *Leisure Sciences, 14*, 317-325. https://doi.org/10.1080/01490409209513177

Welsh, B., Sullivan, C., & Olds, D. (2010). When early crime prevention goes to scale: A new look at the evidence. *Prevention Science, 11*(2), 115-125. https://doi.org/10.1007/s11121-009-0159-4

Wessels, I. (2017). *Parent engagement in parent training interventions: Findings from the Sinovuyo Careing Families Project* (unpublished doctoal dissertation). University of Cape Town. Cape Town, South Africa.

Wessels, I., Lester, S., & Ward, C. (2016). Engagement in parenting programmes: Exploring facilitators of and barriers to participation. http://137.158.155.94/bitstream/handle/11427/19880/PolicyBrief82.pdf?sequence=1

Whittaker, K., & Cowley, S. (2012). An effective programme is not enough: A review of factors associated with poor attendance and engagement with parenting support programmes. *Children & Society*, *26*(2), 138-149. https://doi.org/10.1111/j.1099-0860.2010.00333.x

World Health Organization [WHO]. (2016). *INSPIRE: Seven strategies for ending violence agaisnt children*. http://apps.who.int/iris/bitstream/handle/10665/207717/9789241565356-eng.pdf?sequence=1

WHO. (2018). *INSPIRE handbook: Action for implementing the seven strategies for ending violence against children*. World Health Organization. http://www.who.int/iris/handle/10665/272996.

Wilson, D., & Lipsey, M. (2001). The role of method in treatment effectiveness research: Evidence from meta-analysis. *Psychological Methods*, *6*(4), 413-429. https://doi.org/10.1037/1082-989X.6.4.413

WHO. (2014). *Global Status Report on Violence Prevention 2014*. https://www.who.int/violence_injury_prevention/violence/status_report/2014/en/

Wu, C. (2016). *Growth mixture cace modeling: A statistical approach for evaluating causal effects in longitudinal interventions with noncompliance* (unpublished doctoal dissertation). The University of Nebraska-Lincoln. United States of America.

Yau, L., & Little, R. (2001). Inference for the complier-average causal effect from longitudinal data subject to noncompliance and missing data, with application to a job

training assessment for the unemployed. *Journal of the American Statistical Association, 96*(456), 1232-1244. https://doi.org/10.1198/016214501753381887

Yun, S. H. (2011). Factor structure and reliability of the Revised Conflict Tactics Scales' (CTS2) 10-factor model in a community-based female sample. *Journal of Interpersonal Violence, 26*(4), 719-744. https://doi.org/10.1177/0886260510365857

Zeileis, A. (2006). Object-oriented computation of sandwich estimators. *Journal of Statistical Software, 16*(9), 1-16. https://doi.org/10.18637/jss.v016.i09

Zhang, Z. (2016). Model building strategy for logistic regression: Purposeful selection. *Annals of translational medicine, 4*(6), 111. https://doi.org/10.21037/atm.2016.02.15

Appendices

Appendix A: University of Oxford Approval, Thailand

Oxford Tropical Research Ethics Committee

University of Oxford
Research Services, University Offices
Wellington Square, Oxford OX1 2JD
Tel. +44 (0) 1865 282585
E-mail: oxtrec@admin.ox.ac.uk

To whom it may concern 6 June 2018

Dear Sir/Madam

Title: Randomized controlled trial of an evidence-informed parenting intervention to prevent violence against young children by parents and primary caregivers in Thailand

OxTREC Ref: 8-18

The above study has been designed by Professor Frances Gardner and colleagues at the University of Oxford. It is funded by UNICEF Thailand. I confirm that the University will accept responsibility for the overall initiation, management and/or financing of the study, subject to all local ethical and regulatory requirements being met.

The University maintains Clinical Trial insurance that is appropriate for local requirements in the territories in which the study will be conducted; to the extent that a claim is made against the University, this will apply to this study.

Sponsorship is confirmed subject to the condition that should any substantial amendments or reports be submitted to the Local Ethics Committee, these should also be copied to OxTREC.

Yours faithfully

Signature Removed

Dr Rebecca Bryant
Research Ethics Manager (OxTREC)

Appendix B: Ministry of Public Health Approval, Thailand

Document No. 27 /2018

The Ethical Review Committee for Research in Human Subjects

Ministry of Public Health, Thailand

Title of Project :	Feasibility study of parenting intervention to prevent violence against young children by parents and primary caregivers in Thailand: Phase One.
Protocol Number :	Ref. no. 10/2561
Principle Investigator :	Asst. Professor Sombat Tapanya
Place of proposed study :	Bangkok (Save the Children Thailand Country Office) Nonthaburi (Ministry of Public Health) Udon Thani (Udon Thani hospitals, local community centres, or government offices (Primary Educational Service Area Office Region 1, Nakhon Udon Thani Municipality Office, Udon Thani Provincial Public Health Office, as well as District Public Health Offices and Health Promotion Hospitals)
Project Time line :	1 year
Biomedical sample for future use/ Time :	not collect

Document Approved :

1. Full Protocol in Thai and English Version 4.0 dated 17 Aug 2018
2. Appendix A: Guide for individual interviews with officials, experts and practitioners in Thai and English Version 4.0 dated 17 Aug 2018
3. Appendix B: Participation information sheet and consent form for officials, experts and practitioners in Thai and English Version 4.0 dated 17 Aug 2018
4. Appendix C: Guide for working group in Thai and English Version 4.0 dated 17 Aug 2018
5. Appendix D: Participation information sheet and consent form for members of the Parenting Experts Working Group in Thai and English Version 4.0 dated 17 Aug 2018
6. Appendix E: Demographic Survey in Thai and English Version 4.0 dated 17 Aug 2018
7. Curriculum vitae

Document No. 97 /2018

We also confirm that we are an ethics committee constituted in agreement and in accordance with the ICH-GCP

The Ethical Review Committee for Research in Human Subjects Ministry of Public Health, Thailand had reviewed mainly Thai protocol. In ethical concern, the committee has reviewed and approved for implementation of the research study as above mention, therefore the Thai protocol will be mainly conduct. The protocol must be approved by continuation review for the duration of one year until expired.

Signature Removed

..... Chairman

(Mr. Somsak Akksilp)

Signature Removed

..... Secretary

(Mr. Attasit Srisubat)

Date of Approval 22 August 2018 Date of Expired 21 August 2019

Appendix C: University of Oxford Approval, Philippines

SOCIAL SCIENCES & HUMANITIES
INTER-DIVISIONAL RESEARCH ETHICS COMMITTEE

Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD
Tel: +44(0)1865 616576 Fax: +44(0)1865 280467
ethics@socsci.ox.ac.uk

27 June 2017

Dr Jamie Lachman
Social Policy & Intervention

Dear Dr Lachman,

Research Ethics Approval (CUREC 2)
Ref No: R52091/RE001
Title: Parenting for Lifelong-Philippines (PLH-Philippines) Evaluation Study

The above application has been considered on behalf of the Social Sciences and Humanities Inter-divisional Research Ethics Committee (IDREC) in accordance with the procedures laid down by the University for ethical approval of all research involving human participants.

I am pleased to inform you that, on the basis of the information provided to the IDREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly approval has been granted.

Should there be any subsequent changes to the project, which raise ethical issues not covered in the original application, you should submit details to the IDREC for consideration.

Please note that you may be required to submit an annual progress report on each anniversary of study approval, until the study is completed, and that your study may be selected for review by the SSH IDREC during an annual audit. External funders may also request ethics audits of any studies they have funded.

We have noted that Oxford University is not the lead in this study, which means that ultimate responsibility lies with the lead University.

Yours sincerely,

Signature Removed

Claudia Kozeny-Pelling
Research Ethics Manager and Secretary SSH IDREC

cc: ethics@spi.ox.ac.uk; Prof Frances Gardener

Appendix D: Ateneo De Manila University Approval, Philippines

ATENEO DE MANILA UNIVERSITY
UNIVERSITY RESEARCH ETHICS OFFICE

05 June 2017

Liane Peña Alampay, PhD
Department of Psychology
School of Social Sciences
Ateneo de Manila University

Re: Application for Ethics Approval

Dear Dr. Alampay,

Peace!

We are pleased to inform you that your application for ethics approval for your project entitled: "Parenting for Lifelong Health (PLH) – Masayang Pamilya (MaPa) Evaluation Study" has been reviewed and given clearance to proceed by the University Research Ethics Committee. This approval is based on the protocol you initially submitted on 27 April 2017. You are expected to comply with the policies of the Ateneo de Manila University's Code of Research Ethics and the Philippine Health Research Ethics Board.

This approval is valid until 05 June 2018. Substantial changes made to your protocol must be reported to and approved by the UREC *prior* to implementation, via the Protocol Amendment Form, unless such a change is necessary to avoid immediate harm to the participants. In such a case, the change or amendment must still be reported. Likewise, any unanticipated problems and/or adverse events that involve risks to participants must be reported to the UREC via the Unanticipated Problems Report Form, within one week of the investigator becoming aware of this problem or adverse event. Such reports must be submitted to the University Research Ethics Office (UREO) and succeeding actions of the researcher must be done after consultation with the UREO or UREC. Kindly contact the UREO for questions or more information.

We wish you success in your research undertaking.

Respectfully yours

Signature Removed

Jacklyn R. Cleofas, PhD
Chair
University Research Ethics Committee

Rm 204, 2/F Xavier Hall, Katipunan Avenue, Loyola Heights, 1108 Quezon City, Philippines
Tel +63 2 426 6001 ext 4030, Email univresearchethics@ateneo.edu

Appendix E: Summary of Outcome Evaluation Tools used in Thailand

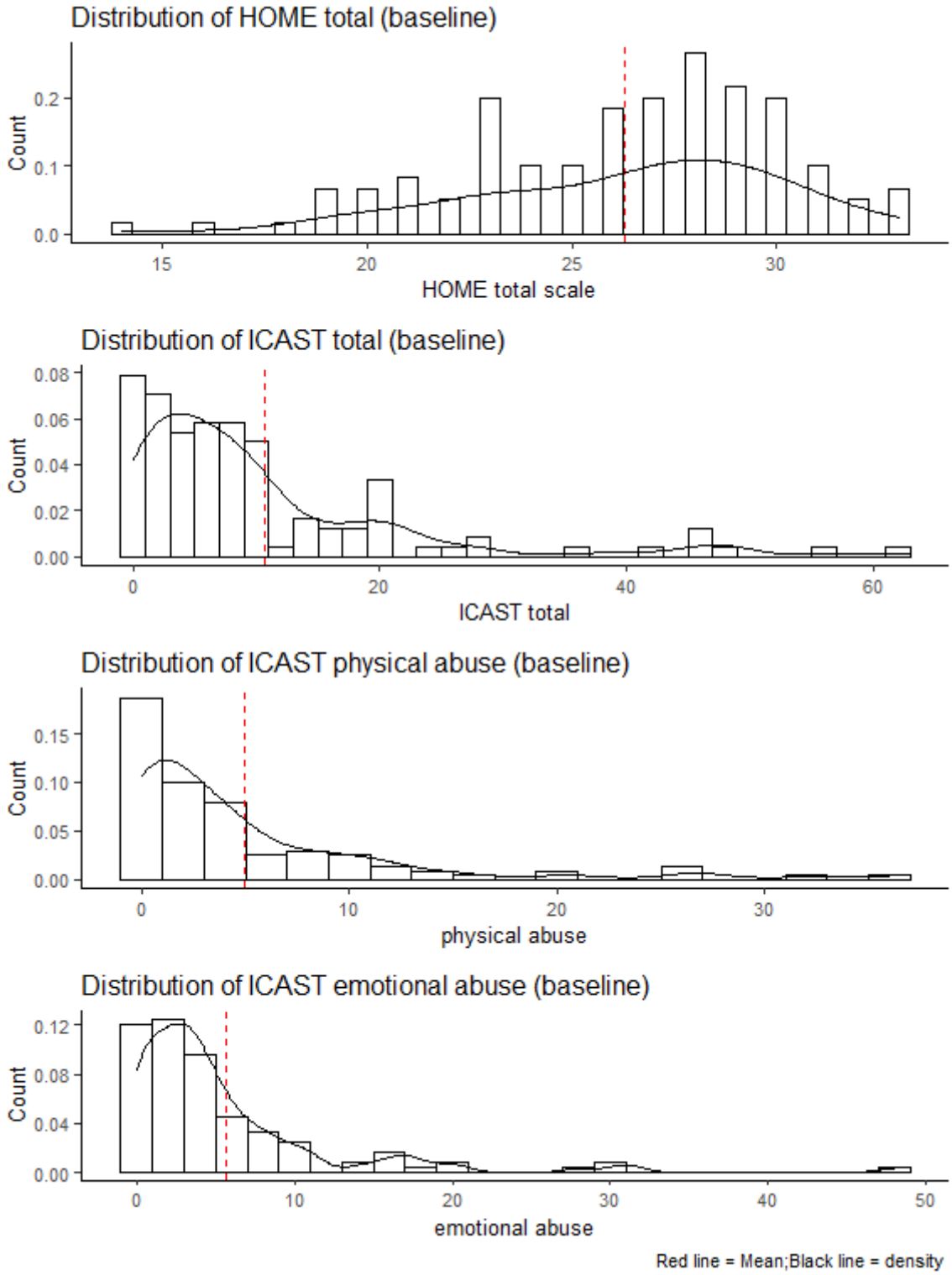
Outcome	Measurement	Timing	Items
Household and family factors			**33**
Parent/primary caregiver, child and family demographics	General demographic questions; Multiple Indicator Cluster Survey	Baseline	24
Parent/primary caregiver physical health	Medical Outcomes Study Short Form – 12 Health Survey	Baseline	3
Parent/primary caregiver and child disability	Washington Group questions (adapted)	Baseline	2
Parent/primary caregiver history of physical abuse, emotional abuse or neglect during childhood	ISPCAN Child Abuse Screening Tool-Retrospective Version (ICAST-R) (adapted)	Baseline	4
Socioeconomic factors			**16**
Household structure	Multiple Indicator Cluster Survey	Baseline	3
Household employment	Multiple Indicator Cluster Survey	Baseline	2
Household assets	Multiple Indicator Cluster Survey	Baseline	2
Household hunger	The Hunger Scale Questionnaire	Baseline	5
Income and benefits	Multiple Indicator Cluster Survey	Baseline	2
	Categories from 2015 National Statistical Office Household Socioeconomic Survey	Baseline	1
Health care coverage	Categories from 2015 National Statistical Office Survey on Health and Welfare	Baseline	1
Primary outcome			**26**
Child maltreatment: physical abuse and emotional abuse	ISPCAN Child Abuse Screening Tool-Trial Caregiver Version (ICAST-T): physical abuse and emotional abuse subscales	Baseline; Post-int.; 3 months follow-up	**20**
	HOME Inventory (Interview & Observational assessment)	Baseline; Post-int.; 3 months follow-up	6
Secondary outcomes			**184**
Positive parenting	Parenting Young Children Scale (PARYC)	Baseline; Post-int.; 3 months follow-up	21
Dysfunctional parenting	Arnold Parenting Scale (PS): Over-reactivity Subscale	Baseline; Post-int.; 3 months follow-up	10
Child monitoring and supervision	Alabama Parenting Questionnaire: Monitoring & supervision subscale	Baseline; Post-int.; 3 months follow-up	11
Child neglect	IPSCAN Child Abuse Screening Tool – Trial Caregiver Version (ICAST-T): neglect subscale	Baseline; Post-int.; 3 months follow-up	6
Parental depression, anxiety and stress	Depression, Anxiety, and Stress Scale (DASS)	Baseline; Post-int.; 3 months follow-up	21
Child behaviour problems	Eyberg Child Behavior Inventory (ECBI)	Baseline; Post-int.; 3 months follow-up	36
Attitudes toward punishment	MICS Child Discipline item	Baseline; Post-int.; 3 months follow-up	1
Parent daily report on child behaviour and parenting	Parent Daily Report Checklist (PDR)	Baseline; 2 time points post-baseline; Post-int.; 3 months follow-up	44

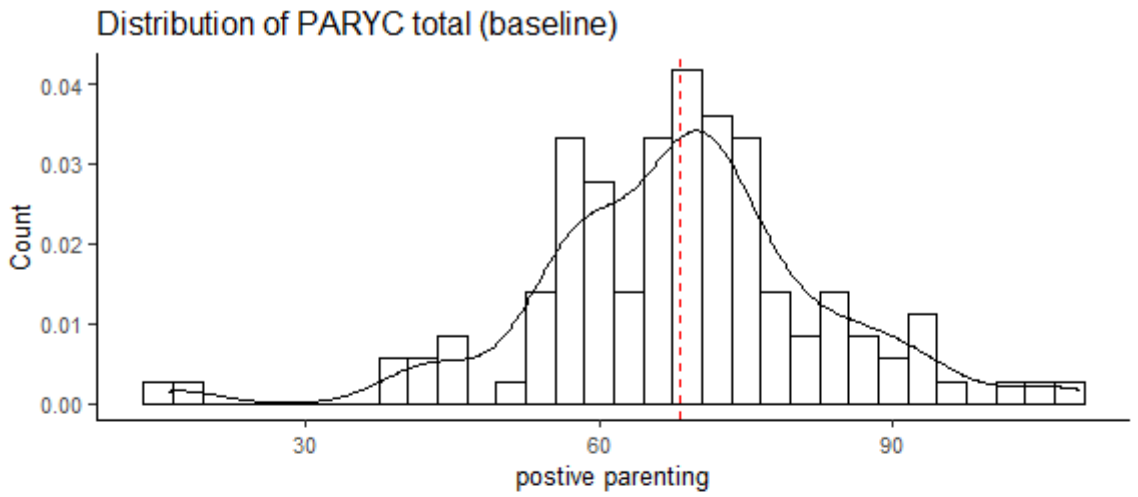
Parent-child relationships	HOME Inventory (Interview & Observational)	Baseline; Post-int.; 3 months follow-up	26
Intimate partner violence	Revised Conflict Tactics Scale Short Form (CTS2S)	Baseline; Post-int.; 3 months follow-up	8
Programme engagement			**15**
Obstacles to programme engagement	Obstacles to Engagement Scale (OES)	Baseline	15

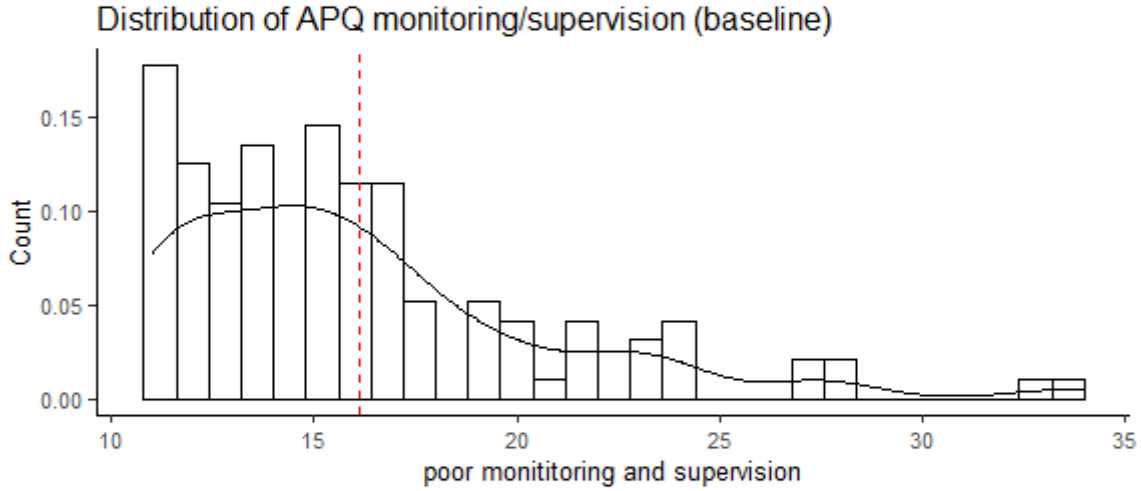
Appendix F: Summary of Outcome Evaluation Tools used in the Philippines

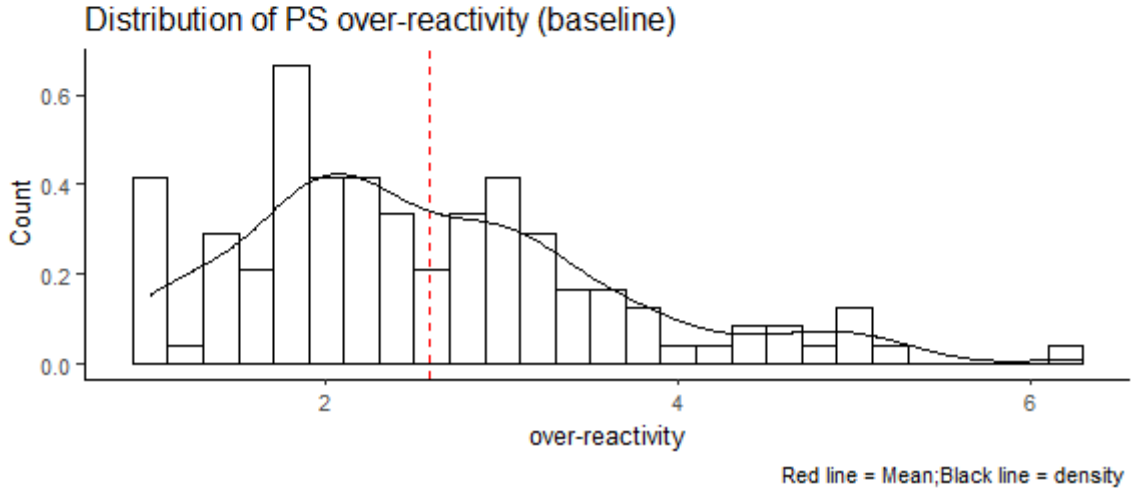
Outcome	Measurement	Items
Demographic factors (potential subgroups and moderators)		**27**
Parent, child and family demographics	General demographics questions	15
Household poverty	The Hunger Scale Questionnaire	3
Parent history of child maltreatment	ISPCAN Child Abuse Screening Tool-Retrospective Version	5
Parent general heath	Medical Outcomes Study Short Form – 12 Health Survey	3
Parent alcohol abuse	Self-report on alcohol use	1
Primary outcome		**22**
Child maltreatment	ISPCAN Child Abuse Screening Tool-Intervention	18
Proximal outcomes		**58**
Positive parenting	Parenting of Young Children Scale	21
Dysfunctional parenting	The Parenting Scale	30
Attitudes to corporal punishment	Multiple Indicator Cluster Survey	1
Daily parenting behaviour	Parent Daily Report Checklist	6
Secondary outcomes (potential mediators)		**181**
Child behaviour problems	Eyberg Child Behaviour Inventory	36
	Parent Daily Report Checklist	35
Child development	Ages and Stages Questionnaire-3rd Edition-Communication Subscale	5
	Ages and Stages Questionnaire-Socio-Emotional	30
Parenting efficacy	Parenting Sense of Competence, Efficacy Subscale	8
Daily parenting efficacy	Parent Daily Report Checklist	3
Parenting stress	Parenting Stress Index	24
Parental mental health	Depression, Anxiety, and Stress Scale	21
	WHO-5 Well-Being Index	5
Parent exposure to IPV	Revised Conflict Tactics Scale Short Form	8
Marital satisfaction	Kansas Marital Satisfaction Survey	3
Parent dependency on alcohol	Alcohol consumption in the past month	1
Parent/child sleep	Average daily hours of sleep in the past 5 days	2
Process evaluation outcomes		**91**
Implementation fidelity	Facilitator check-lists	n/a
	Parenting for Lifelong Health Facilitator Assessment Tool (PLH-FAT)	51
Adherence/exposure/engagement	Attendance registration	n/a
	Activity check-list (parent report)	n/a
Satisfaction/acceptability	Parent overall satisfaction	40
Acceptability and feasibility	Parent in-depth interviews	n/a
	Facilitator focus groups	n/a

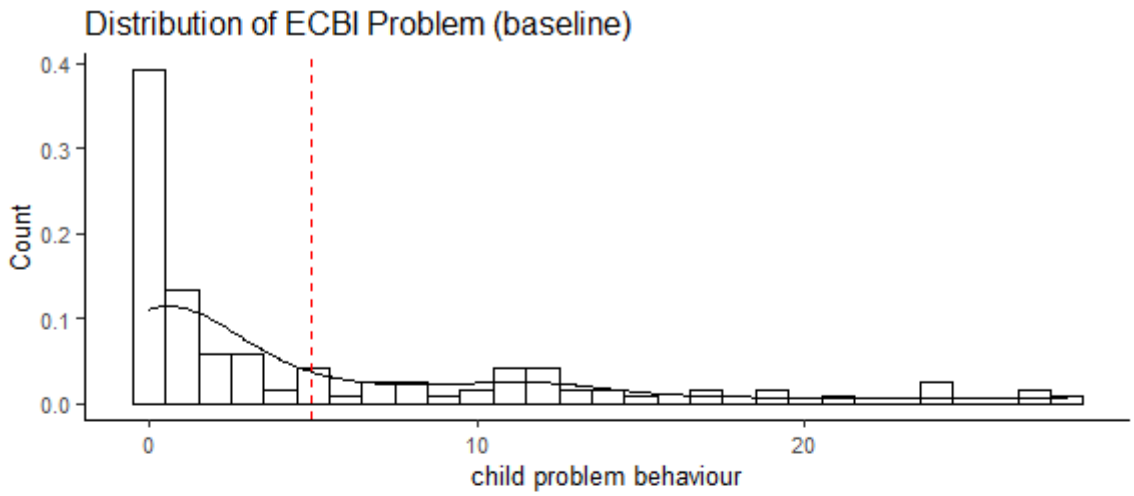
Appendix G: Histograms of Primary and Secondary Outcomes: Thailand

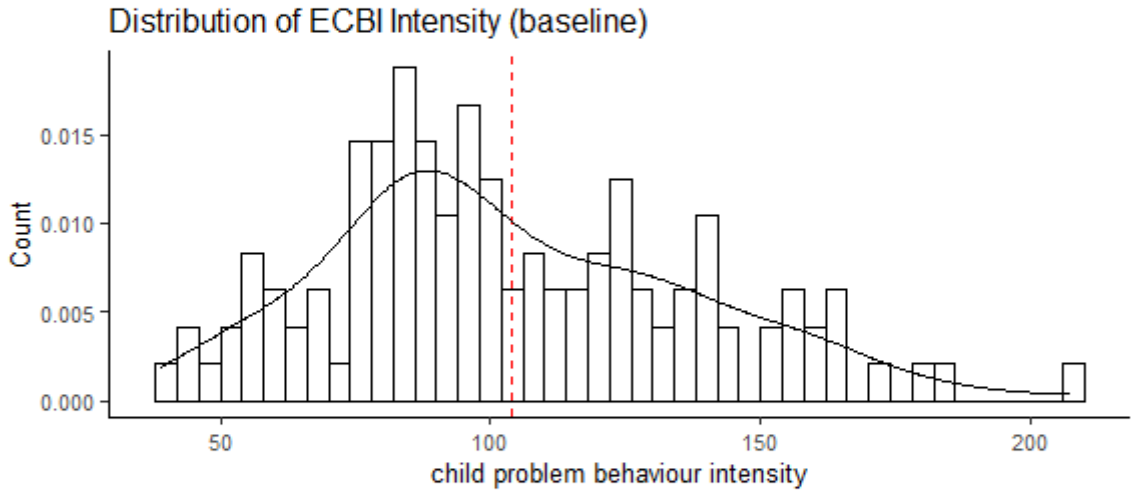


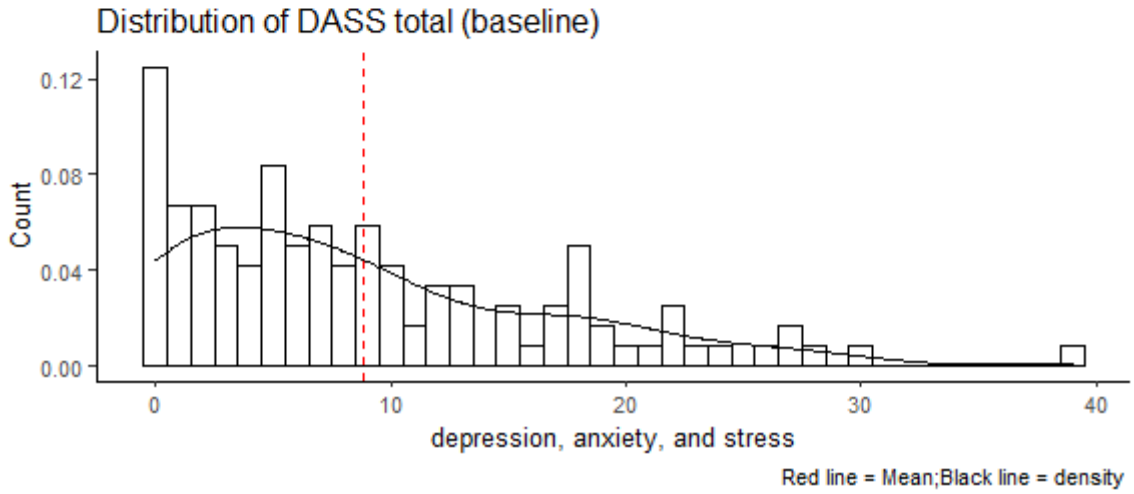

Distribution of PARYC total (baseline)
Count
0.04
0.03
0.02
0.01
0.00
30
60
90
postive parenting


Distribution of APQ monitoring/supervision (baseline)
Count
0.15
0.10
0.05
0.00
10
15
20
25
30
35
poor monititoring and supervision

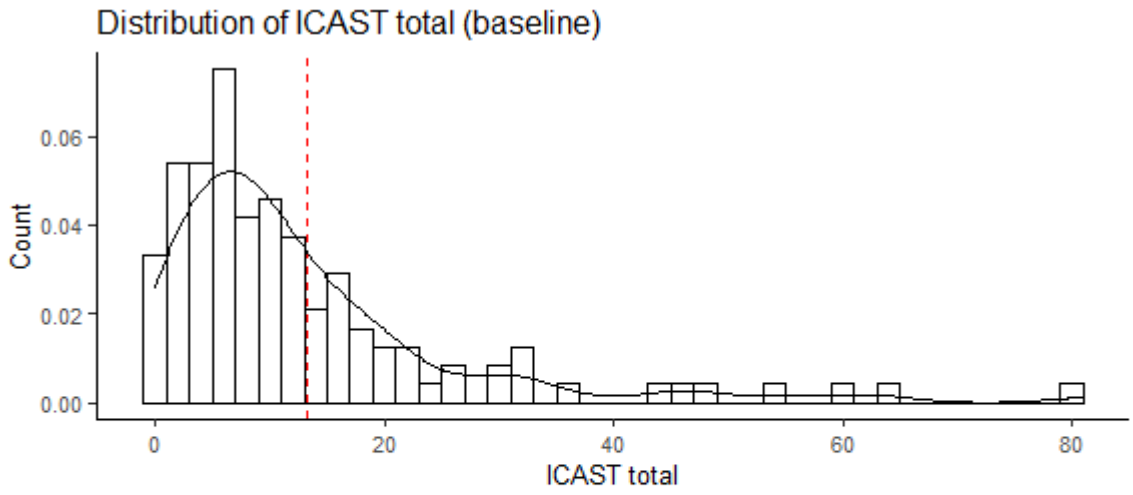

Distribution of PS over-reactivity (baseline)
Count
0.6
0.4
0.2
0.0
2
4
6
over-reactivity
Red line = Mean;Black line = density

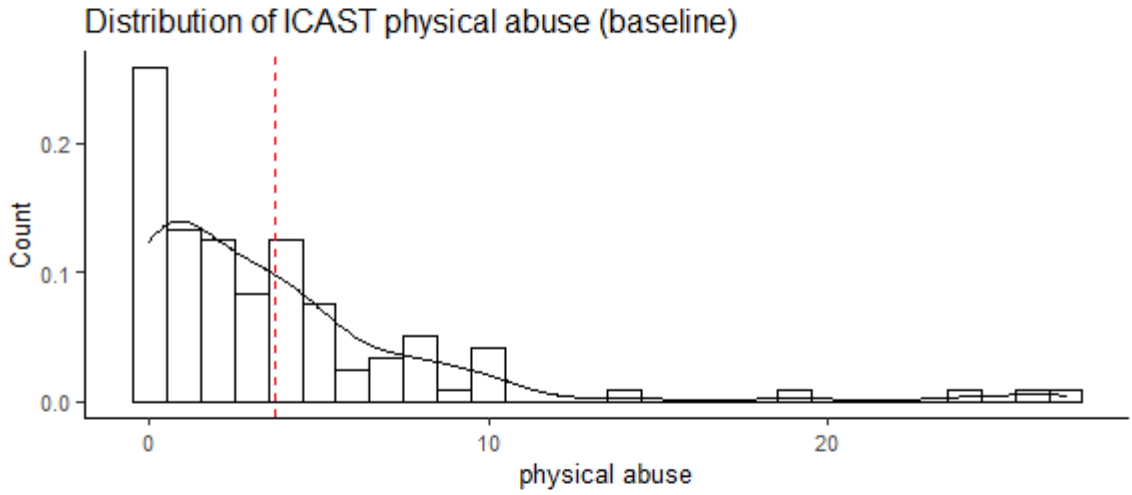

Distribution of ECBI Problem (baseline)
Count
0.4
0.3
0.2
0.1
0.0
0
10
20
child problem behaviour

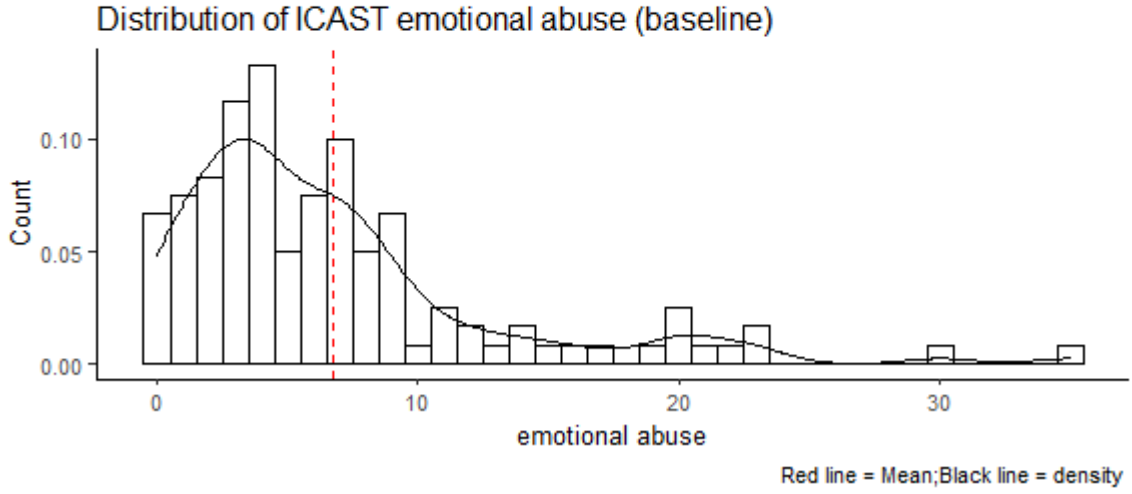

Distribution of ECBI Intensity (baseline)
Count
0.015
0.010
0.005
0.000
50
100
150
200
child problem behaviour intensity


Distribution of DASS total (baseline)
Count
0.12
0.08
0.04
0.00
0
10
20
30
40
depression, anxiety, and stress
Red line = Mean;Black line = density

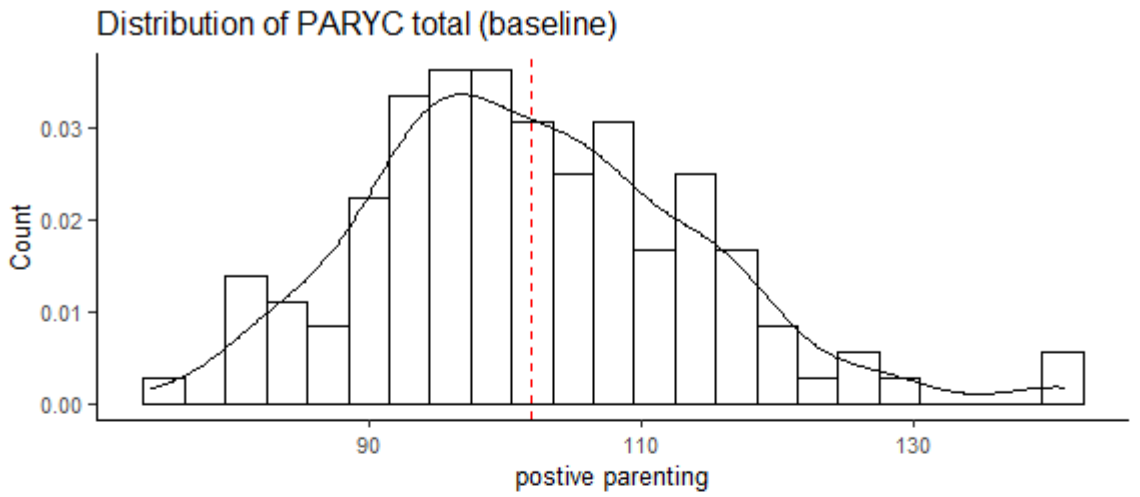
Appendix H: Histograms of Primary and Secondary Outcomes: Philippines

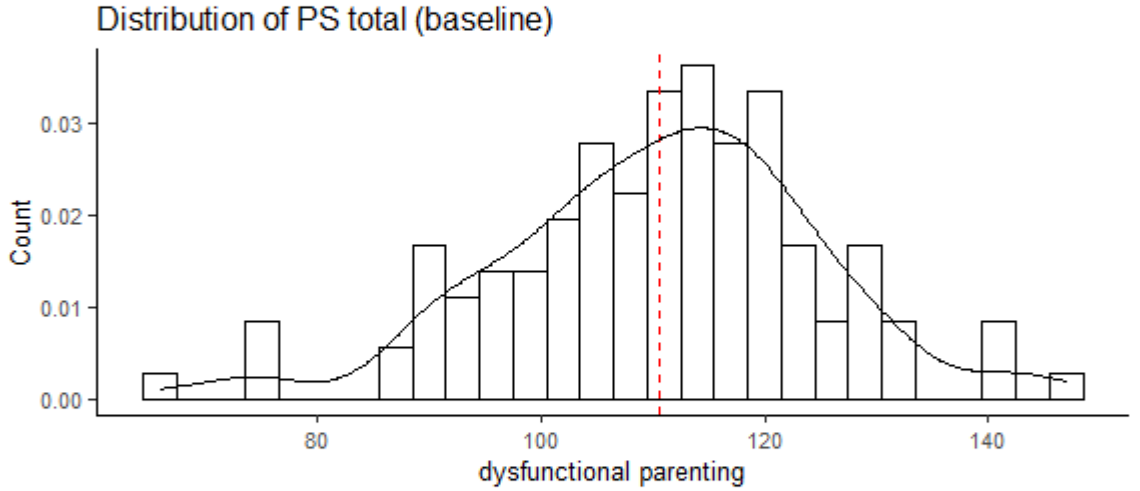


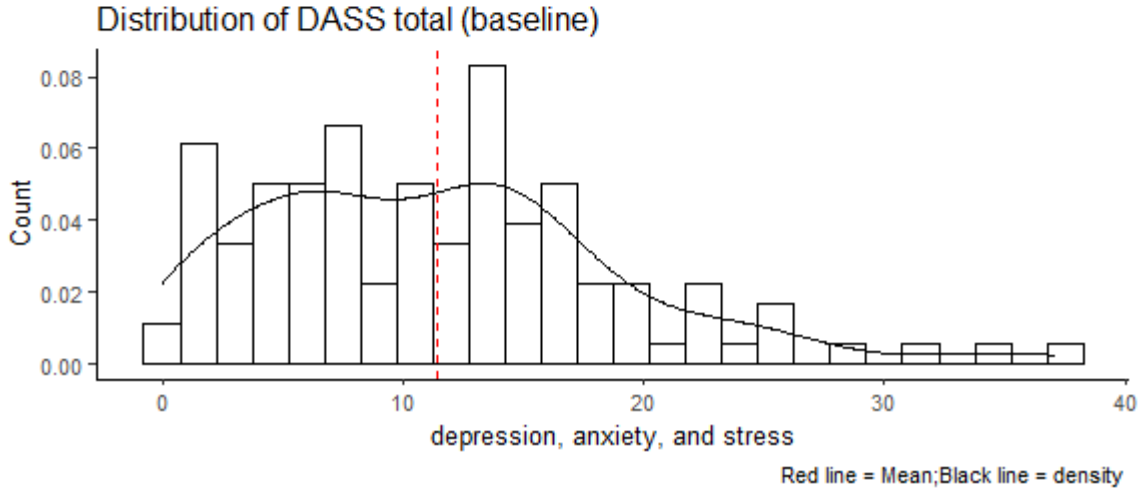


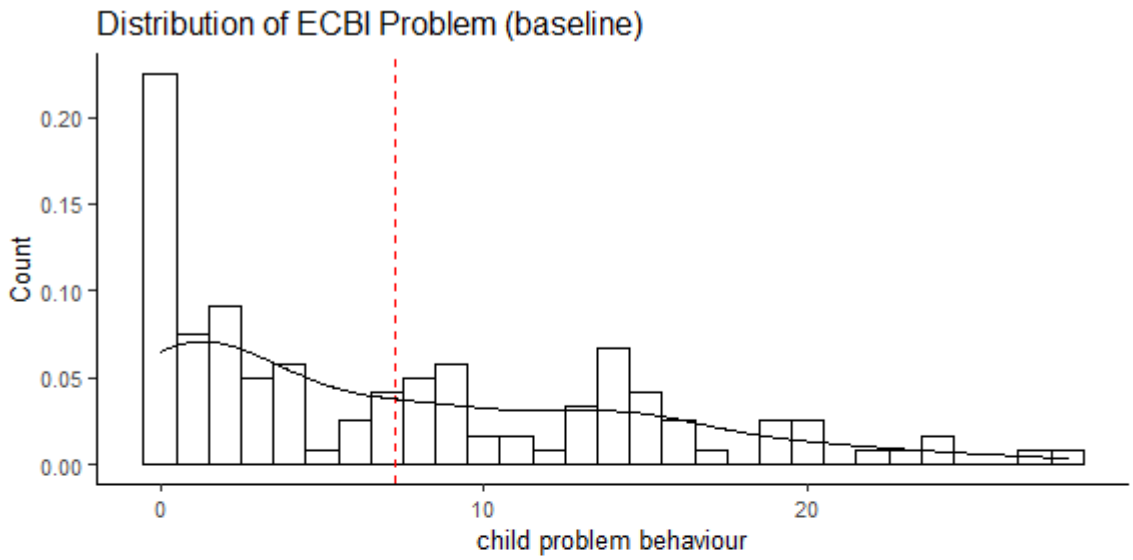


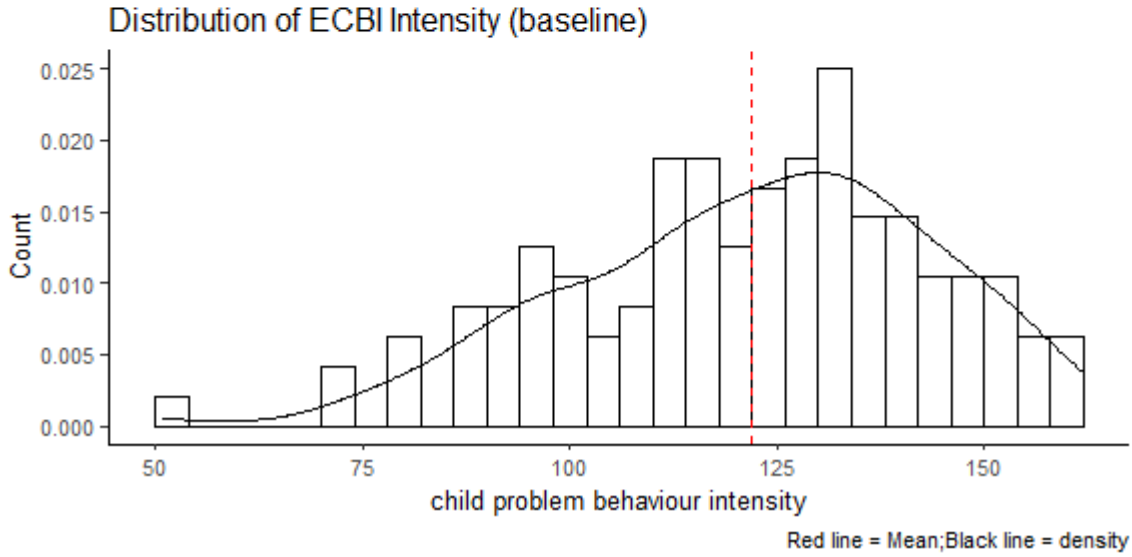
Red line = Mean;Black line = density


Distribution of PARYC total (baseline)
Count
0.03
0.02
0.01
0.00
90
110
130
postive parenting


Distribution of PS total (baseline)
Count
0.03
0.02
0.01
0.00
80
100
120
140
dysfunctional parenting


Distribution of DASS total (baseline)
Count
0.08
0.06
0.04
0.02
0.00
0
10
20
30
40
depression, anxiety, and stress
Red line = Mean;Black line = density

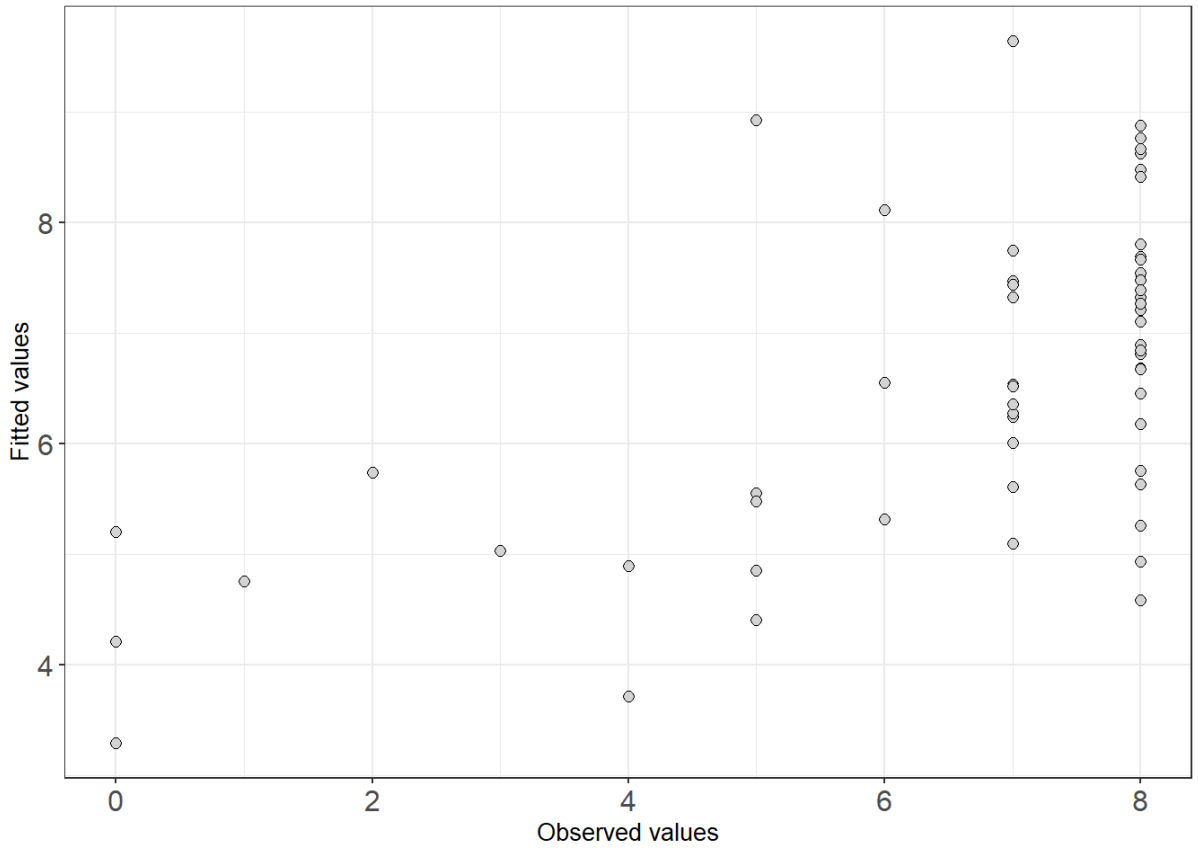

Distribution of ECBI Problem (baseline)
Count
0.20
0.15
0.10
0.05
0.00
0
10
20
child problem behaviour


Distribution of ECBI Intensity (baseline)
Count
0.025
0.020
0.015
0.010
0.005
0.000
50
75
100
125
150
child problem behaviour intensity
Red line = Mean;Black line = density

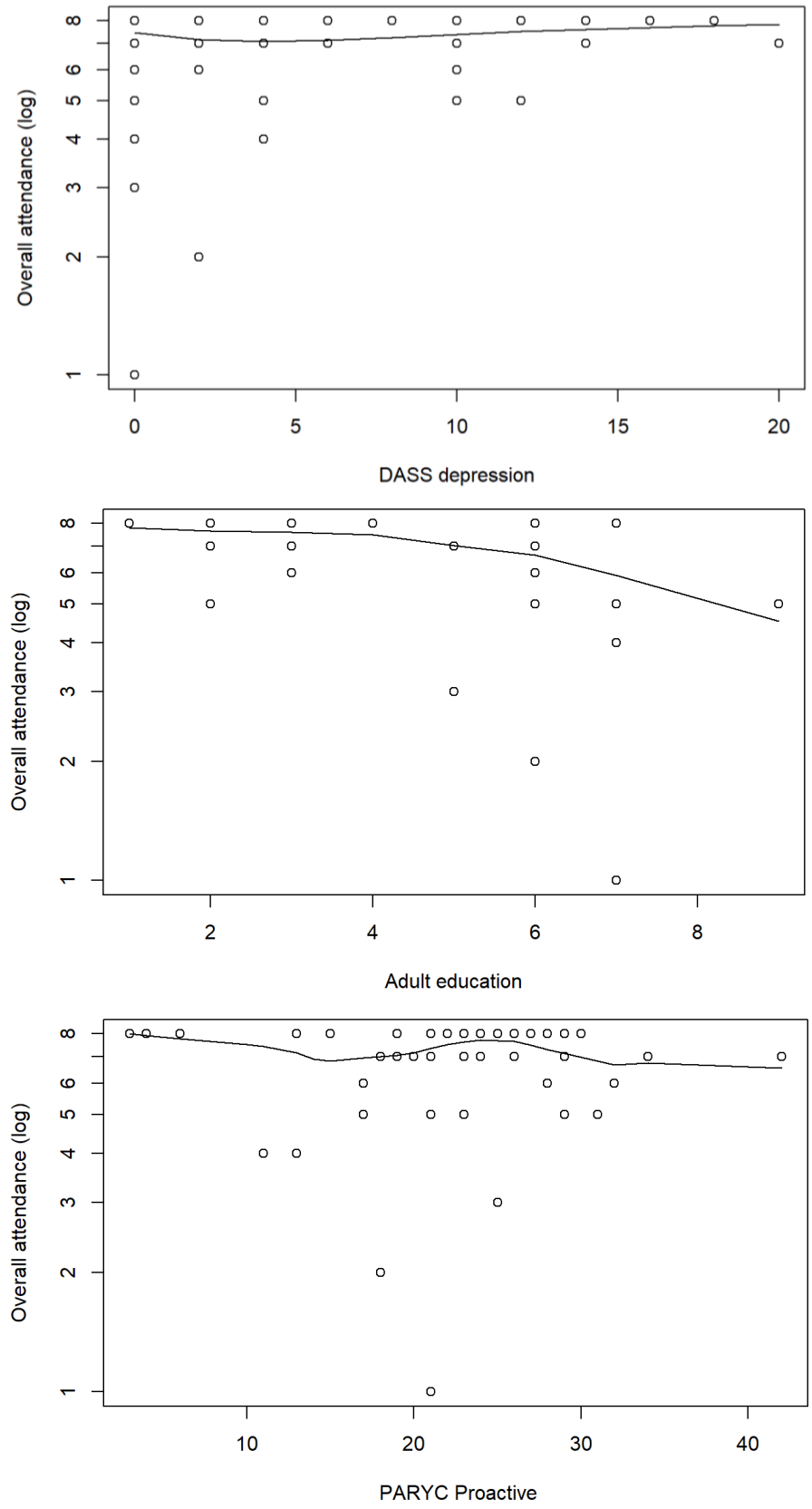
Appendix I: Plotting Fitted Values Against Observed Values: Thailand

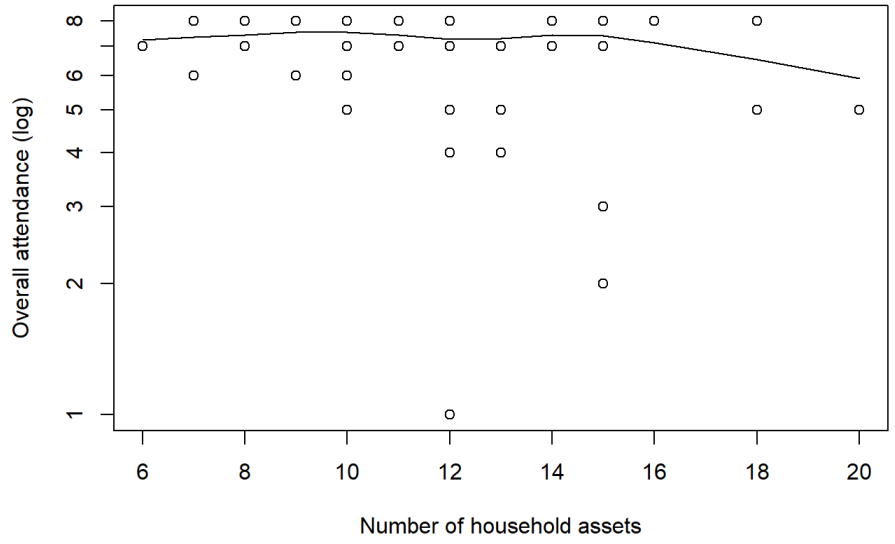
Figure 5

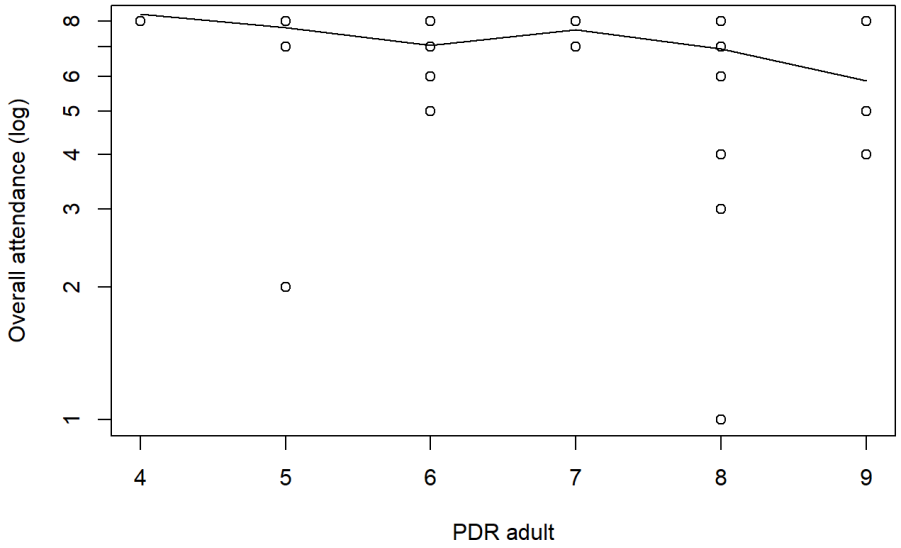
Multivariate Model Fitted Values Vs Observed Values: Thailand



Appendix J: Checking Linearity of Continuous Predictors: Thailand



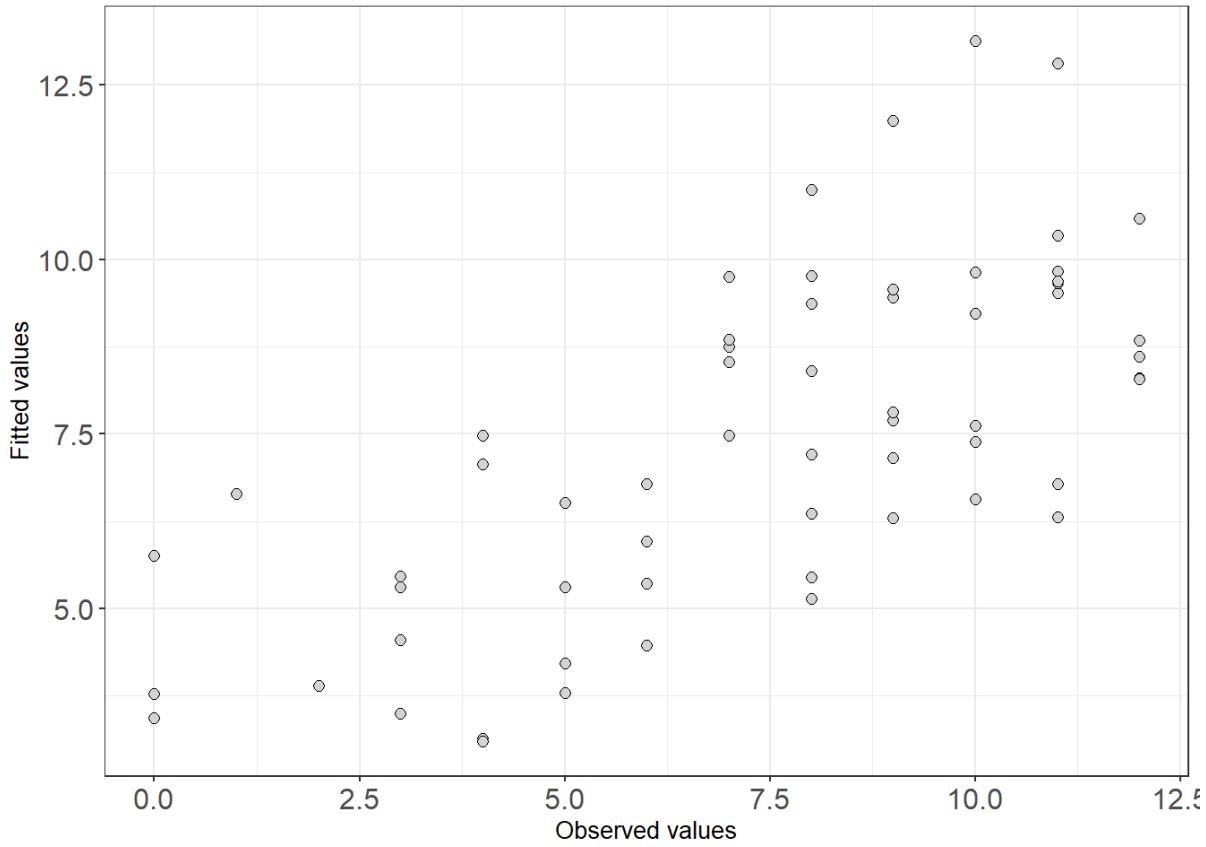

Overall attendance (log)
1
2
3
4
5
6
8
6
8
10
12
14
16
18
20
Number of household assets


Overall attendance (log)
1
2
3
4
5
6
8
4
5
6
7
8
9
PDR adult

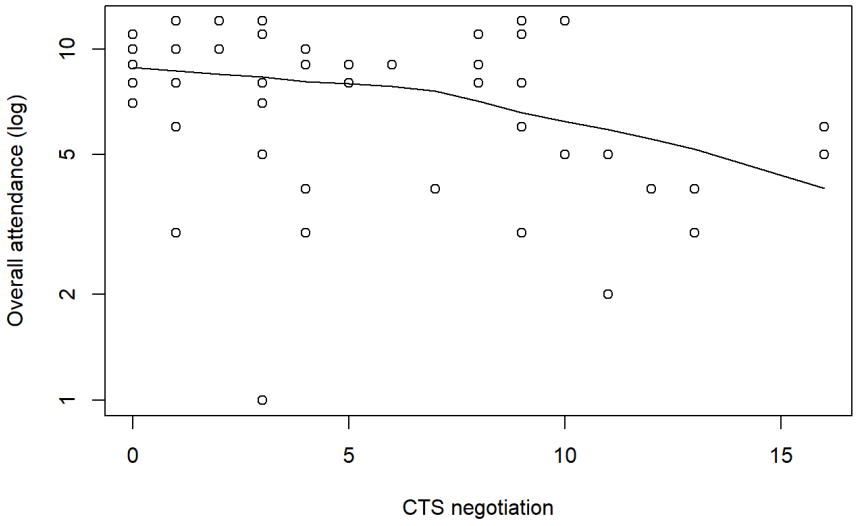
Appendix K: Plotting Fitted Values Against Observed Values: Philippines

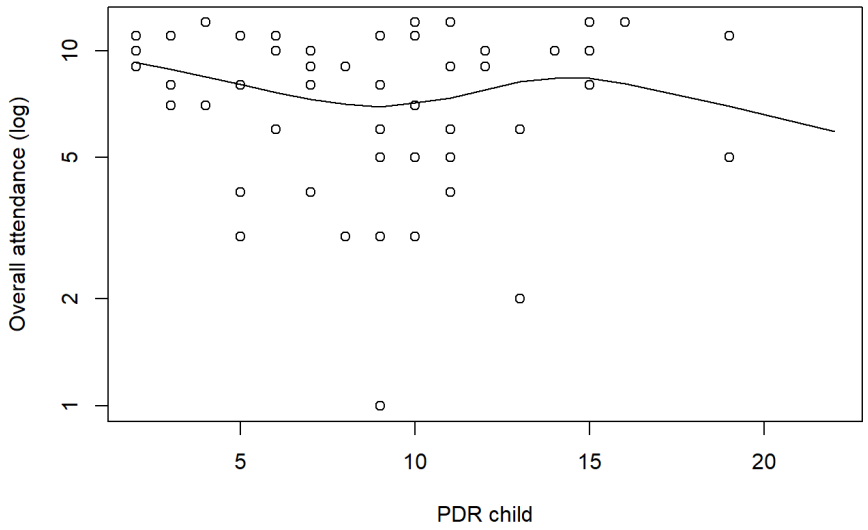
Figure 6

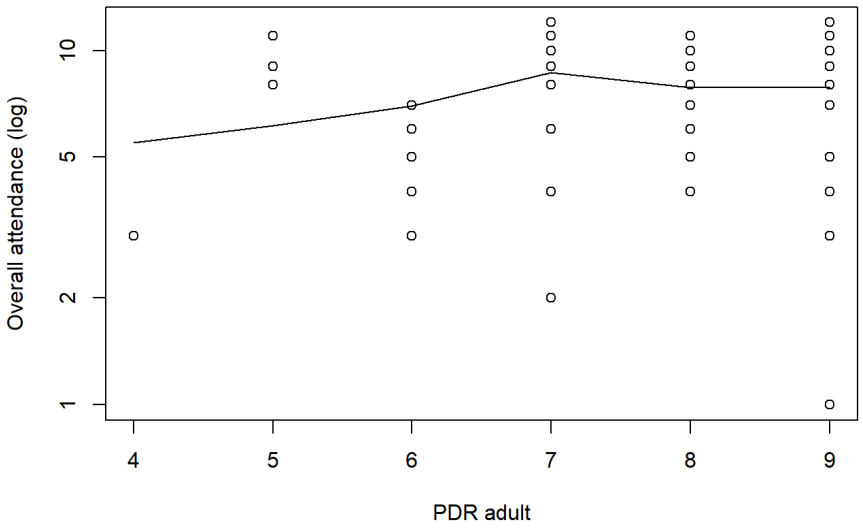
Multivariate Model Fitted Values Vs Observed Values: Philippines

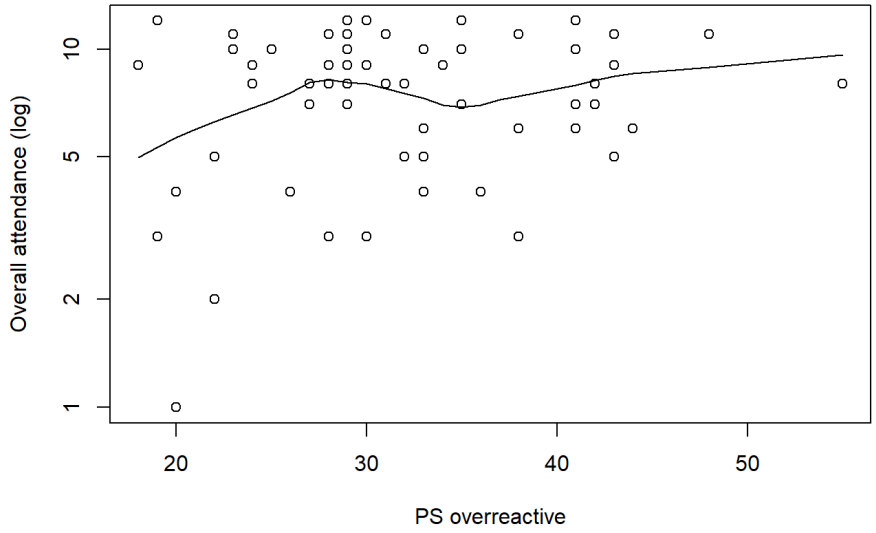


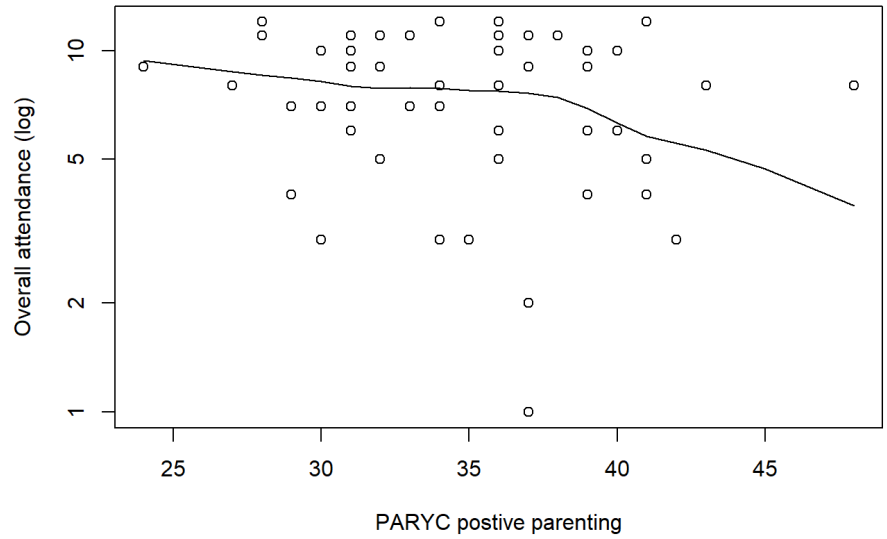
Appendix L: Checking Linearity of Continuous Predictors: Philippines



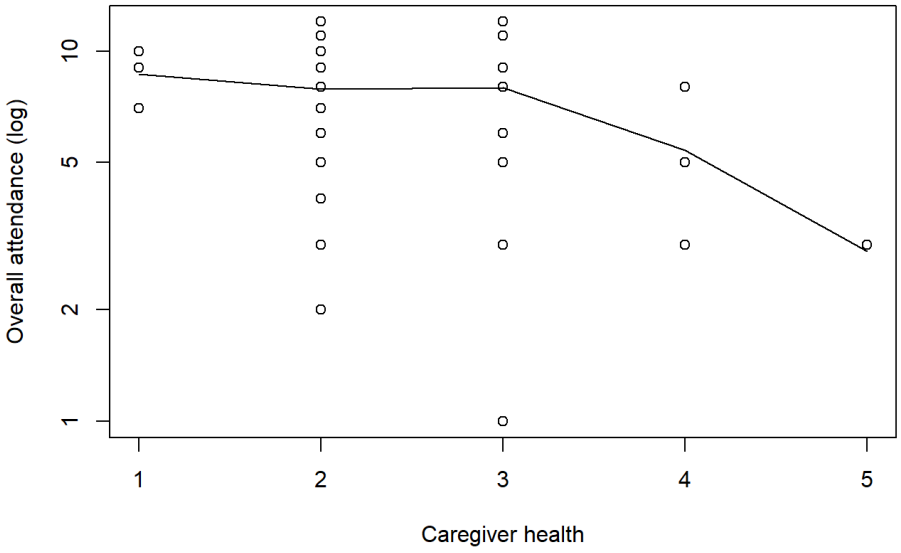


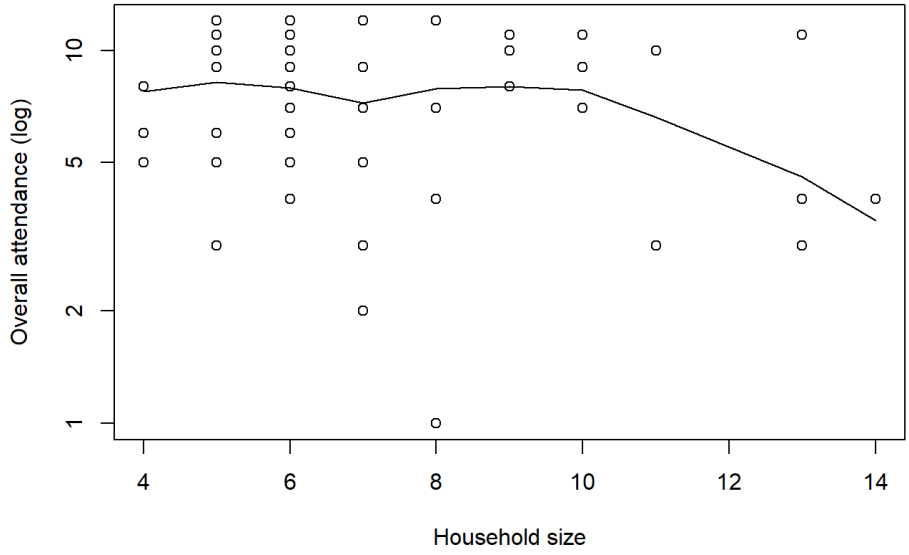


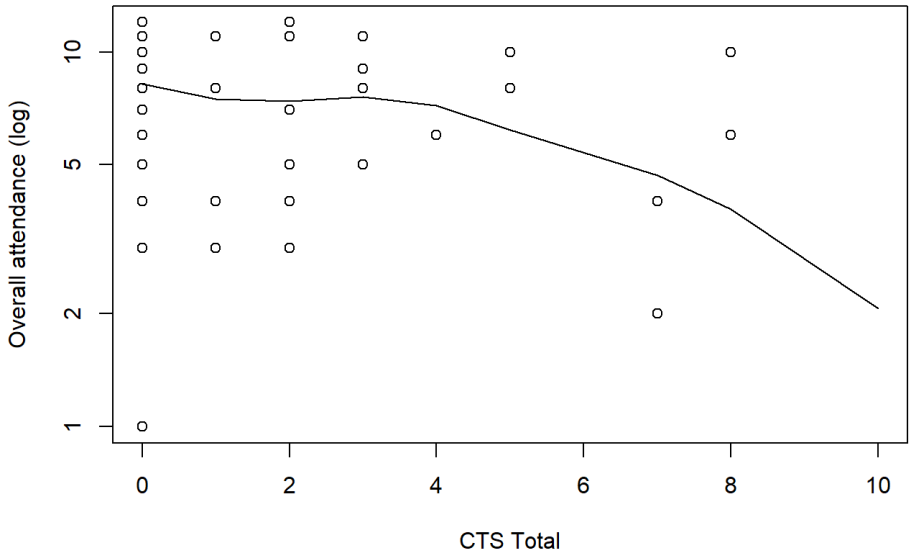

Overall attendance (log)
1
2
5
10
20
30
40
50
PS overreactive


Overall attendance (log)
1
2
5
10
25
30
35
40
45
PARYC postive parenting


Overall attendance (log)
1
2
5
10
0
5
10
15
20
ICAST emotional abuse


Overall attendance (log)
1
2
5
10
1
2
3
4
5
Caregiver health


Overall attendance (log)
1
2
5
10
4
6
8
10
12
14
Household size


Overall attendance (log)
1
2
5
10
0
2
4
6
8
10
CTS Total

Appendix M: Adjusted CACE Models Including Covariates: Thailand

Table 20

CACE Models with Covariates for High Attenders: Thailand

Outcome	Standardised ß [95% CI]	Unstandardised ß [95%CI]	S.E.	*p*	IRR [95% CI]
Harsh parenting [a]					
Posttest	-0.28 [-0.53, -0.04]	-0.70 [-1.23, -0.17]	0.27	.009	0.50 [0.26, 0.90]
Follow-up	-0.31 [-0.56, -0.07]	-0.78 [-1.29, -0.26]	0.26	.003	0.46 [0.25, 0.83]
Child maltreatment (overall) [a]					
Posttest	-0.51 [-0.60, -0.42]	-1.27 [-1.69, -0.42]	0.30	<.001	0.28 [0.22, 0.35]
Follow-up	-0.42 [-0.52, -0.32]	-1.05 [-1.54, -0.54]	0.32	.001	0.35 [0.27, 0.45]
Physical child abuse [a]					
Posttest	-0.53 [-0.67, -0.39]	-1.32 [-2.01, -0.63]	0.35	<.001	0.27 [0.19, 0.38]
Follow-up	-0.46 [-0.64, -0.29]	-1.15 [-2.01, -0.22]	0.47	.015	0.31 [0.20, 0.47]
Emotional child abuse [a]					
Posttest	-0.51 [-0.63, -0.38]	-1.26 [-1.93, -0.60]	0.34	<.001	0.28 [0.21, 0.38]
Follow-up	-0.39 [-0.52, -0.26]	-0.97 [-1.64, -0.31]	0.34	.004	0.38 [0.27, 0.51]
Home (total) [b]					
Posttest	0.30 [0.08, 0.68]	2.11 [0.50, 3.73]	0.82	.010	
Follow-up	0.23 [0.10, 0.55]	2.00 [0.34, 3.67]	0.85	.020	
Positive parenting (total) [b]					
Posttest	0.64 [0.26, 1.02]	10.40 [4.31, 16.49]	3.11	<.001	
Follow-up	0.64 [0.23, 1.04]	9.64 [3.73, 15.55]	3.02	.001	
Positive parenting [b]					
Posttest	0.89 [0.48, 1.30]	4.34 [2.40, 6.28]	0.99	<.001	
Follow-up	0.60 [0.18, 1.03]	3.01 [0.98, 5.04]	1.04	.004	
Setting limits [b]					
Posttest	0.81 [0.39, 1.23]	5.60 [2.67, 8.52]	1.49	<.001	
Follow-up	0.61 [0.18, 1.05]	3.65 [1.06, 6.25]	1.32	.012	
Proactive parenting [b]					
Posttest	0.17 [-0.24, 0.59]	1.18 [-1.61, 3.96]	1.42	.408	
Follow-up	0.64 [0.23, 1.04]	3.75 [1.54, 5.96]	1.12	<.001	
Overreactive parenting [b]					
Posttest	-0.60 [-1.05, -0.15]	-0.57 [-0.99, -0.16]	0.21	.007	
Follow-up	-0.47 [-0.92, -0.01]	-0.43 [-0.84, -0.03]	0.21	.037	
APQ monitoring [b]					
Posttest	-0.68 [-1.04, -0.31]	-3.17 [-4.87, -1.47]	0.87	<.001	
Follow-up	-0.45 [-0.84, -0.07]	-2.22 [-3.94, -0.49]	0.88	.012	
Child problem behaviour [a]					
Posttest	-0.55 [-0.68, -0.41]	-1.36 [-2.18, -0.54]	0.42	.001	0.26 [0.18, 0.36]

Outcome	Standardised ß [95% CI]	Unstandardised ß [95%CI]	S.E.	*p*	IRR [95% CI]
Follow-up	-0.63 [-0.78, -0.49]	-1.58 [-2.41, -0.75]	0.42	<.001	0.21 [0.14, 0.30]
Child problem behaviour intensity [b]					
Posttest	-0.66 [-0.97, -0.35]	-23.39 [-33.70, -13.09]	5.26	<.001	
Follow-up	-0.51 [-0.83, -0.19]	-17.61 [-28.08, -7.14]	5.34	<.001	
DASS (total) [a]					
Posttest	-0.25 [-0.32, -0.17]	-0.61 [-1.03, -0.19]	0.22	.004	0.54 [0.45, 0.66]
Follow-up	-0.23 [-0.31, -0.14]	-0.56 [-1.00, -0.12]	0.22	.012	0.57 [0.46, 0.70]
Depression [a]					
Posttest	-0.41 [-0.51, -0.30]	-1.00 [-1.58, -0.42]	0.30	<.001	0.37 [0.28, 0.47]
Follow-up	-0.40 [-0.51, -0.28]	-0.99 [-1.56, -0.24]	0.29	<.001	0.37 [0.28, 0.49]
Anxiety [a]					
Posttest	-0.19 [-0.29, -0.09]	-0.47 [-0.94, 0.01]	0.24	.053	0.63 [0.49, 0.80]
Follow-up	-0.22 [-0.32, -0.11]	-0.55 [-1.02, -0.07]	0.24	.025	0.58 [0.45, 0.75]
Stress [a]					
Posttest	-0.20 [-0.28, -0.11]	-0.49 [-0.92, -0.06]	0.22	.027	0.61 [0.49, 0.76]
Follow-up	-0.15 [-0.24, -0.05]	-0.37 [-0.87, 0.13]	0.25	.144	0.69 [0.55, 0.87]

Note. [a]Poisson logistic regression analyses; [b]Linear regression analyses; IRR = Incidents rate ratio; SE = Robust sandwich estimators

Appendix N: Adjusted CACE Models Including Covariates: Philippines

Table 21

CACE Models with Covariates for High Attenders: Philippines

Outcome	Standardised ß [95%CI]	Unstandardised ß [95%CI]	S.E.	*p*	IRR [95% CI]
Child maltreatment (overall) [a]					
Posttest	-0.35 [-0.42, -0.29]	-1.36 [-2.00, -0.84]	0.27	<.001	0.26 [0.19, 0.33]
Follow-up	-0.23 [-0.30, -0.17]	-1.00 [-1.50, -0.30]	0.31	.003	0.41 [0.31, 0.53]
Physical child abuse [a]					
Posttest	-0.36 [-0.51, -0.21]	-1.40 [-2.38, -0.42]	0.50	.005	0.25 [0.14, 0.43]
Follow-up	-0.25 [-0.38, -0.12]	-0.98 [-1.92, -0.04]	0.48	.040	0.37 [0.22, 0.63]
Emotional child abuse [a]					
Posttest	-0.36 [-0.45, -0.27]	-1.37 [-1.98, -0.75]	0.31	<.001	0.26 [0.18, 0.36]
Follow-up	-0.19 [-0.28, -0.10]	-0.71 [-1.36, -0.06]	0.33	.032	0.49 [0.35, 0.69]
Positive parenting (total) [b]					
Posttest	0.59 [-0.16, 1.34]	7.59 [-2.28, 17.47]	5.04	.132	
Follow-up	0.41 [-0.29, 1.11]	5.42 [-3.84, 14.68]	4.72	.251	
Positive parenting [b]					
Posttest	0.58 [-0.17, 1.32]	7.41 [-2.69, 3.91]	5.16	.151	
Follow-up	0.19 [-0.57, 0.94]	2.47 [-7.44, 12.38]	5.06	.625	
Monitoring [b]					
Posttest	0.36 [-0.38, 1.10]	1.70 [-1.69, 5.09]	1.73	.326	
Follow-up	0.56 [-0.18, 1.31]	2.98 [-0.81, 6.77]	1.93	.123	
Proactive parenting [b]					
Posttest	0.43 [-0.36, 1.23]	2.36 [-1.96, 6.68]	2.20	.285	
Follow-up	-0.06 [-0.80, 0.68]	-0.36 [-4.39. 3.68]	2.06	.863	
Dysfunctional parenting [b]					
Posttest	-1.74 [-2.49, -0.98]	-20.97 [-30.03, -11.90]	4.62	<.001	
Follow-up	-0.37 [-1.08, 0.34]	-5.52 [-15.91, 4.87]	5.30	.298	
Laxness [b]					
Posttest	-0.67 [-1.39, 0.05]	-5.08 [-10.57, 0.41]	2.80	.070	
Follow-up	-0.05 [-0.81, 0.70]	-0.51 [-7.52, 6.50]	3.58	.888	
Overreactive [b]					
Posttest	-1.78 [-2.60, -0.97]	-12.8 [-18.69, -6.95]	2.99	<.001	
Follow-up	-0.64 [-1.32, 0.05]	-4.97 [-10.12, 0.17]	2.63	.058	
Verbosity [b]					
Posttest	-1.30 [-2.19, -0.40]	-6.54 [-10.76, -2.32]	2.15	.002	
Follow-up	0.17 [-0.72, 1.05]	0.89 [-3.85, 5.62]	2.42	.714	
Child problem behaviour [a]					

Outcome	Standardised ß [95%CI]	Unstandardised ß [95%CI]	S.E.	*p*	IRR [95% CI]
Posttest	-0.11 [-0.19, -0.03]	-0.44 [-1.28, 0.40]	0.43	.306	0.65 [0.47, 0.89]
Follow-up	0.07 [-0.53, 0.67]	0.25 [-0.55, 1.06]	0.41	.537	1.29 [0.91. 1.83]
Child problem behaviour intensity [b]					
Posttest	-0.67 [-1.36, 0.01]	-16.48 [-32.69, -0.26]	8.27	.046	
Follow-up	-0.33 [-1.05, 0.39]	-7.76 [-24.91, 9.38]	8.75	.375	
DASS (total) [b]					
Posttest	-0.05 [-0.72, 0.61]	-0.42 [-5.37, 4.53]	2.53	.867	
Follow-up	0.13 [-0.58, 0.85]	1.00 [-4.17, 6.16]	2.63	.705	
Depression [b]					
Posttest	-0.17 [-0.92, 0.58]	-0.96 [-4.78, 2.86]	1.95	.624	
Follow-up	-0.05 [-0.83, 0.33]	-0.26 [-4.00, 3.49]	1.91	.893	
Anxiety [b]					
Posttest	-0.11 [-0.76, 0.54]	-0.67 [-4.42, 3.08]	1.91	.725	
Follow-up	-0.06 [-0.74, 0.62]	-0.37 [-4.51, 3.77]	2.11	.860	
Stress [b]					
Posttest	0.03 [-0.69, 0.74]	0.17 [-4.36, 4.70]	2.31	.941	
Follow-up	0.29 [-0.45, 1.02]	1.82 [-2.66, 6.30]	2.29	.425	

Note. [a]Poisson logistic regression analyses; [b]Linear regression analyses; IRR = Incidents rate ratio; SE = Robust sandwich estimators